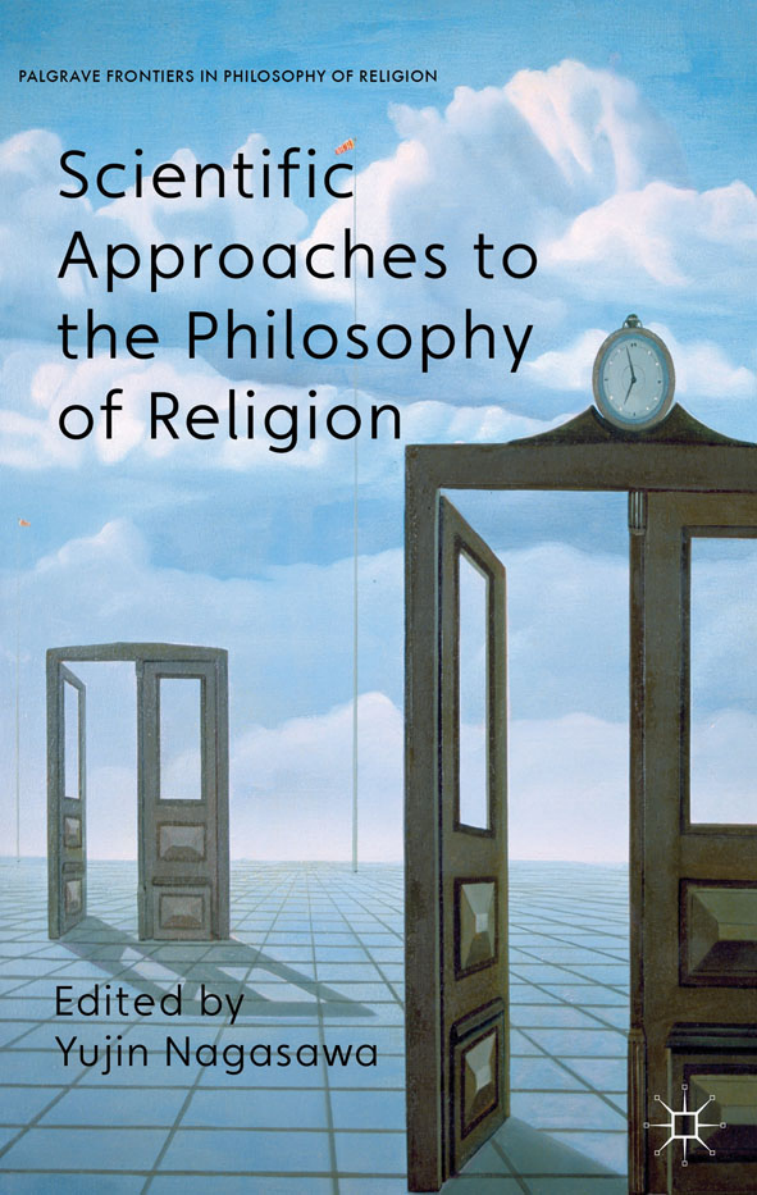
PALGRAVE FRONTIERS IN PHILOSOPHY OF RELIGION

Scientific Approaches to the Philosophy of Religion



Edited by
Yujin Nagasawa

Scientific Approaches to the Philosophy of Religion

Palgrave Frontiers in Philosophy of Religion

Series Editors: Yujin Nagasawa and Erik J. Wielenberg

Titles Include:

Yujin Nagasawa (*editor*)
SCIENTIFIC APPROACHES TO THE PHILOSOPHY OF RELIGION

Forthcoming titles

Gregory Dawes and James Maclaurin (*editors*)
COGNITIVE SCIENCE AND RELIGION

Trent Dougherty
THE PROBLEM OF ANIMAL PAIN:
A Theodicy for All Creatures Great and Small

Aaron Rizzieri
PRAGMATIC ENCROACHMENT, RELIGIOUS BELIEF AND PRACTICE

Also by Yujin Nagasawa

THE EXISTENCE OF GOD:
A Philosophical Introduction

GOD AND PHENOMENAL CONSCIOUSNESS:
A Novel Approach to Knowledge Arguments

NEW WAVES IN PHILOSOPHY OF RELIGION (co-editor with Erik Wielenberg)

THERE'S SOMETHING ABOUT MARY:
Essays on Phenomenal Consciousness and Frank Jackson's Knowledge Argument
(co-editor with Peter Ludlow and Daniel Stoljar)

Scientific Approaches to the Philosophy of Religion

Edited by

Yujin Nagasawa
University of Birmingham, UK

palgrave
macmillan

Selection and editorial matter © Yujin Nagasawa 2012
Chapters © their individual authors 2012

All rights reserved. No reproduction, copy or transmission of this publication may be made without written permission.

No portion of this publication may be reproduced, copied or transmitted save with written permission or in accordance with the provisions of the Copyright, Designs and Patents Act 1988, or under the terms of any licence permitting limited copying issued by the Copyright Licensing Agency, Saffron House, 6–10 Kirby Street, London EC1N 8TS.

Any person who does any unauthorized act in relation to this publication may be liable to criminal prosecution and civil claims for damages.

The authors have asserted their rights to be identified as the authors of this work in accordance with the Copyright, Designs and Patents Act 1988.

First published 2012 by
PALGRAVE MACMILLAN

Palgrave Macmillan in the UK is an imprint of Macmillan Publishers Limited, registered in England, company number 785998, of Houndmills, Basingstoke, Hampshire RG21 6XS.

Palgrave Macmillan in the US is a division of St Martin's Press LLC, 175 Fifth Avenue, New York, NY 10010.

Palgrave Macmillan is the global academic imprint of the above companies and has companies and representatives throughout the world.

Palgrave® and Macmillan® are registered trademarks in the United States, the United Kingdom, Europe and other countries.

ISBN 978–0–230–29110–2

This book is printed on paper suitable for recycling and made from fully managed and sustained forest sources. Logging, pulping and manufacturing processes are expected to conform to the environmental regulations of the country of origin.

A catalogue record for this book is available from the British Library.

A catalog record for this book is available from the Library of Congress.

10 9 8 7 6 5 4 3 2 1
21 20 19 18 17 16 15 14 13 12

Printed and bound in Great Britain by
CPI Antony Rowe, Chippenham and Eastbourne

Contents

Series Preface vii

Acknowledgements viii

Notes on Contributors ix

Introduction 1
Yujin Nagasawa

Part I Divine Attributes

1 The Necessity of God and the Psychology of Counterfactual Thinking 11
Robin Le Poidevin

2 Why Would Anyone Believe in a Timeless God? Two Types of Theology 28
Benjamin Murphy

Part II God, Creation and Evolution

3 Darwin's Argument from Evil 49
Paul Draper

4 Attributing Agency: Fast and Frugal or All Things Considered? 71
Graham Wood

Part III God and the Universe

5 On Non-Singular Space-times and the Beginning of the Universe 95
William Lane Craig and James D. Sinclair

6 The Theistic Multiverse: Problems and Prospects 143
Klaas J. Kraay

Part IV Religious Beliefs

7 How Relevant Is the Cognitive Science of Religion to Philosophy of Religion? 165
David Leech and Aku Visala

8 The Rationality of Classical Theism and Its Demographics 184
T. J. Mawson

Part V Religious Tolerance and Disagreement

9 Coercion, Consequence and Salvation 205
Steve Clarke

10 Polarized Yet Warranted Christian Belief 224
David Efird

Part VI The Compatibility of Science and Religion

11 Freedom, Science and Religion 237
Katherin A. Rogers

12 The Compatibility of Science and Religion: Why the Warfare Thesis Is False 255
Michael Ruse

Index 275

Series Preface

The philosophy of religion has experienced a welcome re-vitalization over the last fifty years or so and is now thriving. Our hope with the *Palgrave Frontiers in Philosophy of Religion* series is to contribute to the continued vitality of the philosophy of religion by producing works that truly break new ground in the field.

Accordingly, each book in this series advances some debate in the philosophy of religion by offering a novel argument to establish a strikingly original thesis or approaching an ongoing dispute from a radically new point of view. Each book accomplishes this by utilising recent developments in empirical sciences or cutting-edge research in foundational areas of philosophy, or by adopting historically neglected approaches.

We expect the series to enrich debates within the philosophy of religion both by expanding the range of positions and arguments on offer and establishing important links between the philosophy of religion and other fields, including not only other areas of philosophy but the empirical sciences as well.

Our ultimate aim, then, is to produce a series of exciting books that explore and expand the frontiers of the philosophy of religion and connect it with other areas of inquiry. We are grateful to Palgrave Macmillan for taking on this project as well as to the authors of the books in the series.

Yujin Nagasawa
Erik J. Wielenberg

Acknowledgements

This volume originated in my project 'Anselmian Perfect-Being Theology and the Cognitive Science of Religion', which was undertaken as part of the Cognition, Religion and Theology Project at Oxford University, funded by the John Templeton Foundation. I would like to thank them for their generous financial and academic support. It should be noted, however, that the views expressed in this volume are not necessarily those of the Cognition, Religion and Theology Project, Oxford University or the John Templeton Foundation. I would also like to thank Priyanka Gibbons, Melanie Blair and other staff of Palgrave Macmillan for their impeccable editorial support. I am also grateful to three anonymous reviewers for helpful comments and constructive suggestions.

Yujin Nagasawa

Notes on Contributors

Steve Clarke is James Martin Fellow in the Institute for Ethics and Science, Oxford Martin School and Faculty of Philosophy, University of Oxford, UK, and Senior Research Fellow in the Centre for Applied Philosophy and Public Ethics, Charles Sturt University, Australia.

William Lane Craig is Research Professor of Philosophy at Talbot School of Theology, USA.

Paul Draper is Professor of Philosophy at Purdue University, USA.

David Efird is Senior Lecturer in Philosophy at the University of York, UK.

Klaas J. Kraay is Associate Professor of Philosophy at Ryerson University, Canada.

Robin Le Poidevin is Professor of Metaphysics at the University of Leeds, UK.

David Leech is Postdoctoral Research Fellow at the Universities of Paris 10 and Blaise Pascal University, France.

T. J. Mawson is Fellow and Tutor in Philosophy at St Peter's College, University of Oxford, UK.

Benjamin Murphy is Associate Professor of Philosophy and Religious Studies at the Florida State University, Panama.

Yujin Nagasawa is Professor of Philosophy of Religion and Co-Director of the John Hick Centre for Philosophy of Religion at the University of Birmingham, UK.

Katherin A. Rogers is Professor of Philosophy at the University of Delaware, USA.

Michael Ruse is the Lucyle T. Werkmeister Professor of Philosophy and Director of the Program in the History and Philosophy of Science at Florida State University, USA.

James D. Sinclair is a physicist and Senior Combat Analyst for the United States Navy, USA.

Aku Visala is Research Fellow at the Academy of Finland and the Ian Ramsey Centre for Science and Religion at the University of Oxford, UK.

Graham Wood is Lecturer in Philosophy at the University of Tasmania, Australia.

Introduction

Yujin Nagasawa

1. The God of Abraham, the God of the Philosophers and the God of the Scientists

The terms the 'God of Abraham' and the 'God of the philosophers' correspond to two traditional approaches to religion. In the first approach, call it the 'supernatural approach', religious believers try to comprehend the existence and nature of God primarily through supernatural means, such as mystical experiences and biblical revelations. In the second approach, call it the 'conceptual approach', philosophers and theologians try to comprehend the existence and nature of God primarily through rational and analytical thinking.

Relatively recently, however, a third approach has emerged. This approach might be called the 'scientific approach' and the concept of God understood through this approach the 'God of the scientists'. On this approach we comprehend the existence and nature of God by appealing to relevant empirical research and scientific studies. This is an interesting combination of the above two traditional approaches. On the one hand, like the supernatural approach, it relies (at least partly) on empirical observations rather than pure conceptual analysis. On the other hand, like the conceptual approach, it adopts rational, analytical thinking. The scientific approach contributes to debates in the philosophy of religion and philosophical theology by addressing relevant research in natural, behavioural and mathematical sciences.

The present volume collects essays that consider central problems in the philosophy of religion and philosophical theology through various forms of the scientific approach (often combined with the conceptual approach). Each chapter aims to advance a debate in the philosophy of religion by referring to specific findings in recent scientific

research. By utilizing theories and discoveries in cognitive science, anthropology, developmental psychology, decision theory, biology, physics and cosmology the present volume tackles several important problems in the philosophy of religion, such as divine attributes; God, creation and evolution; God and the universe; religious beliefs; religious tolerance and disagreement; and the compatibility of science and religion.

2. Divine Attributes

We typically ascribe many impressive attributes – such as omniscience, omnipotence, omnibenevolence, timelessness, incorporeality and necessary existence – to God. But why do we do that? There are two traditional responses to this question. The first such response is that religious experiences and biblical revelations teach us that God manifests these impressive attributes. This response corresponds to the supernatural approach mentioned above. The other response is that God is, by definition, the being than which no greater can be thought and this definition entails that God has these specific attributes. This response corresponds to the conceptual approach mentioned above. Over the last few years, however, researchers in cognitive science, anthropology and developmental psychology have defended an alternative account. According to them, people ascribe these impressive attributes to God because of cognitive and evolutionary facts about human nature. Part I of this volume addresses the debate on divine attributes by referring to these scientific studies.

In Chapter 1 Robin Le Poidevin focuses on God's necessary existence. God's necessary existence is normally interpreted in terms of *de dicto* or *de re* reading of it but each of the interpretations faces difficulties of its own. Le Poidevin, hence, proposes a third interpretation, a projectivist interpretation, of understanding it that is free from these difficulties. He explains, by appealing to the psychology of counterfactual situations, why the assumption of God's necessary existence is such a natural one for theists to make.

Cognitive scientists have studied people's ascription of superlative attributes to specific agents. For example, they have done empirical research on children's ascription of omnipotence and omniscience to adults (Barrett, 2004). In Chapter 2, Benjamin Murphy focuses on another superlative attribute that is commonly ascribed to God, namely, timelessness. Reviewing relevant empirical research he provides two distinct reasons for believing in a timeless God. Murphy argues that

these reasons are rooted in two distinct types of theological thinking, which he calls technical theology and experience-enhancing theology, respectively.

3. God, Creation and Evolution

The dispute over evolution and creation is perhaps the most widely discussed topic concerning religion and science. William Paley published his seminal work *Natural Theology* in 1802, wherein he defended the design argument for the existence of God. This argument purports to derive the existence of God as an intelligent designer by referring to certain features, mainly biological features, of the world. Paley's argument was well received because there was no fully satisfactory naturalistic account of biological diversity and complexity in his time. Thus even Richard Dawkins, one of the fiercest contemporary opponents of creationism, remarks that 'Paley's argument is made with passionate sincerity and is informed by the best biological scholarship of his day' (Dawkins, 1985, p. 5). However, with the publication of Charles Darwin's *On the Origin of Species* in 1859, Paley's theory was no longer tenable. Darwin succeeded in establishing a very powerful naturalistic account of the origin of the biological diversity and complexity. Nevertheless, disputes over evolution among theists and atheists continue.

In response to the evolution–creation debate, some theists have argued that the theory of evolution is compatible with theism. That is, they argue that theists can consistently hold the claim that while the theory of evolution is correct God is still the ultimate creator of the world. However, in Chapter 3, Paul Draper opposes such a claim. He contends that Darwin's theory of evolution motivates a form of the argument from evil against the existence of God. Draper examines Darwin's autobiography and other writings, and maintains that this argument is strongly suggested by Darwin himself. The argument purports to show that the theory of natural selection is more compatible with atheism than it is with theism. Draper strengthens the argument by utilizing Darwin's ideas, his own ideas and findings in evolutionary biology.

In Chapter 4 Graham Wood approaches the notion of design from another perspective. He focuses on the attribution of agency, which seems to be central to religious belief. For example, when theists claim that God created the universe they effectively attribute agency to the cause of the universe. Wood offers, relying on evolutionary psychology and the philosophy of mind, a characterization of the cognitive process that explains how and why we attribute agency to entities, including

religious entities. Using Daniel Dennett's insight on the intentional stance, Wood introduces the attribution of agency in religious contexts, which he calls the 'spiritual stance'.

4. God and the Universe

Among arguments for the existence of God, the design argument has attracted the most attention over the last few years. However, there are many other arguments discussed by philosophers of religion. The cosmological argument is one of them. Some theists claim that contemporary cosmology reinforces the cosmological argument and the credibility of the existence of God as a creator of the universe. Addressing the big bang theory, Pope Pius XII once said, 'true science discovers God in an ever increasing degree – as though God were waiting behind every door' (*Time*, 3 December 1951). Since he made this remark in the 1950s, cosmologists have developed much more advanced theories and hypotheses concerning the origin of the universe.

In Chapter 5 William Lane Craig and James D. Sinclair focus on the *Kalam* cosmological argument. The *Kalam* argument is a version of the cosmological argument originally developed by medieval Islamic philosophers as a conceptual argument for the existence of God as the first cause of the universe. Craig has revived the argument and attempted to strengthen it by appealing to contemporary cosmology. Chapter 5 is a thorough discussion of the cosmological underpinning of the *Kalam* argument. Craig and Sinclair claim that space-time, featuring an initial cosmological singularity, is likely to have a beginning, which is good news for the *Kalam* argument. However, they claim, that does not mean that the argument relies on the truth of the singular nature of space-time. With reference to relevant scientific findings and theories they argue that the *Kalam* argument can be defended even if space-time is non-singular.

In Chapter 6 Klaas J. Kraay addresses the multiverse hypothesis, another currently disputed topic in cosmology and theoretical physics. According to this hypothesis, physical reality consists of multiple universes, including our own universe. Kraay relates scientists' discussion of the multiverse hypothesis with the idea in the philosophy of religion concerning God's actualization of possible universes. He assesses the claim that multiverse theories can help theists in responding to the problem of evil and concludes that it reframes the problem of evil, rather than strengthening theists' or atheists' position on the problem of evil.

5. Religious Beliefs

One of the most important questions in the philosophy of religion, at least in the Judaeo–Christian–Islamic tradition, is the existence of God. Philosophers of religion have developed a number of arguments for and against the existence of God over the last few hundreds of years. Since the 1960s, however, questions concerning the *justification* of, as opposed to *arguments for and against*, theism seem to have attracted more attention. That is, philosophers of religion have intensified their interest in considering whether theists are justified in believing that God exists, rather than whether they can demonstrate that God exists. In recent years cognitive scientists and psychologists have conducted empirical studies concerning the naturalistic origin of beliefs in God and supernatural entities. Based on such studies some have concluded that people, especially small children, are 'intuitive theists' because they are naturally predisposed to hold beliefs about supernatural nonhuman designers (Kelemen, 2004). A challenge for theists here is to justify theistic beliefs in spite of naturalistic explanations of their origins.

In Chapter 7, David Leech and Aku Visala focus on the cognitive science of religion. Atheists often appeal to findings in the cognitive science of religion to undermine theism while theists try to show the contrary – some theists even claim that the cognitive science of religion motivates theism. Leech and Visala examine some implications of the cognitive science in this context. They argue that it is reasonable to conclude that the cognitive science of religion supports neither theism nor atheism and, thus, it should not make any significant impact on the dispute between theists and atheists.

In Chapter 8, T. J. Mawson focuses on the demographics of theism, that is, the geographical and cultural distribution of the population that subscribes to theism. He first raises some worries about the methods used to determine the distribution. He then considers its implications to theism. He examines a version of the argument from evil that appeals to the demographics of theism. Roughly speaking, the argument says that the actual, uneven distribution of belief in God spread throughout the world differs from what one would expect were a perfect God to exist. Mawson explains why this argument fails and concludes that theists must see the demographics of theism not as a reason to think that we do not know that God exists, let alone as a reason to think that God does not exist.

6. Religious Tolerance and Disagreement

Religion has been a source of conflict throughout human history. As society grows more diverse, religious conflicts appear to multiply and intensify. Inter-religious relations and religious diversity have thus attracted significant attention in the philosophy of religion in recent years. Religious exclusivists hold that their own religions are correct and that all others are incorrect. Religious inclusivists hold that while their own religions are correct and privileged they are compatible with at least some others. Religious pluralists hold that their own religions are correct and that, moreover, there are other religions that are equally correct. The cogency of these views has been a matter of controversy among philosophers and theologians for a long time. Recent research in psychology and decision theory has provided useful tools with which to address problems that arise from religious diversity.

In Chapter 9, Steve Clarke focuses on a specific form of exclusivism called salvific exclusivism. A typical form of salvific exclusivism says that those who fail to accept a particular set of religious beliefs are excluded from the possibility of salvation. This view has been influential throughout the history of Christianity and Islam. Salvific exclusivism is dangerous from the point of view of liberal societies if it is combined with a commitment to the moral goal of maximizing the best consequences for others. Clarke models salvific exclusivism, in particular the version he calls interventionist salvific exclusivism, by utilizing contemporary work on decision-making under risk and considers what the government of the liberal state can do to influence leaders of salvific exclusivism.

In Chapter 10 David Efird applies psychological research on belief polarization to religious disagreement. Belief polarization is a phenomenon in which two people with conflicting views tend to become more, rather than less, confident in their views when presented with mixed evidence. In such a situation people harden, rather than soften, their disagreements. It appears that belief polarization undermines the warrant of religious belief as it seems to show that religious belief is not the product of cognitive processes aimed at the truth. Efird, however, provides a model of belief with which he aims to show that at least Christian belief can be polarized yet warranted.

7. The Compatibility of Science and Religion

Chapters 1 through 10 are concerned with specific scientific findings and their implications for specific issues in the philosophy of religion.

However, there is a broader issue concerning the compatibility of science and religion. Some claim that science and religion are fundamentally incompatible, so those who take science seriously cannot hold religious beliefs. Yet others have argued against such a view. The final two chapters of this volume address this issue.

In Chapter 11, Katherin A. Rogers addresses research on free will in experimental psychology. First, she expresses her concerns about a practical danger in denying human freedom by appealing to evidence found in current research. She suggests ways in which future experiments could be conducted and evaluated so as to avoid some of the worrisome aspects. She predicts that, from a Christian perspective, while science may uncover limitations regarding our ability to choose, it will not show that we are unfree in a philosophically and theologically significant way.

In Chapter 12 Michael Ruse considers the compatibility of science and religion from an even more general point of view. He examines the so-called 'warfare' thesis, according to which science and religion are in conflict and that if we hold to the one, we have to give up the other. He argues against the warfare thesis and defends instead the 'independence' thesis, according to which science and religion are independent from each other and hence that they are not in conflict. He argues that what science can grasp is necessarily limited and that it is legitimate for us to try to answer questions reaching beyond the grasp of science.

References

Barrett, J. L. (2004) *Why Would Anyone Believe in God?* (Lanham, MD: Altamira Press).

Darwin, C. (1998 [1859]) *On the Origin of Species* (Oxford: Oxford University Press).

Dawkins, R. (1985) *The Blind Watchmaker* (London: W.W. Norton).

Kelemen, D. (2004) 'Are Children "Intuitive Theists"? Reasoning about Purpose and Design in Nature', *Psychological Science* 15: 295–301.

Paley, W. (2006 [1802]) *Natural Theology* (Oxford: Oxford University Press).

This page intentionally left blank

Part I
Divine Attributes

This page intentionally left blank

1

The Necessity of God and the Psychology of Counterfactual Thinking

Robin Le Poidevin

> And certainly this being so truly exists that it cannot be even thought not to exist.
>
> – Anselm, *Proslogion*

1. Introduction

That God is not a merely contingent or accidental being, one whose birth and death are a matter of chance, is part of what, in orthodox theistic thinking, separates him from us. His existence, unlike ours, is not determined by anything external. While we wither and perish, he is incorruptible. Moreover, independence and incorruptibility are themselves not accidental properties of God, but essential ones. These aspects of his being, suggests John Findlay (as a prelude to an atheistic proof), are required if God is to be truly worthy of worship:

> The true object of religious reverence must not be one, merely, to which no *actual* independent realities stand opposed: it must be one to which such opposition is totally *inconceivable*. God mustn't merely cover the territory of the actual, but also, with equal comprehensiveness, the territory of the possible. And not only must the existence of *other* things be unthinkable without him, but his own non-existence must be wholly unthinkable in any circumstances. (Findlay, 1955, p. 52; emphasis original)

The necessity of God's existence also offers to play an explanatory role, as indicated by a familiar modal form of the cosmological argument: All contingent beings have a cause of their existence, but to avoid the infinite regress this would otherwise threaten, and a lack of

explanation for the existence of the series of contingent causes as a whole, we must posit a first cause whose existence is not contingent, but necessary. This argument may not convince, of course, but even if we do not agree that we are obliged to posit a necessary cause, its role in the argument provides a reason for thinking that *if* God exists he does so of necessity. But do we really understand what it would be for a being to be necessary? The proposition that an object has such-and-such a property as a matter of necessity can be interpreted as *de dicto* or *de re*. That there are *de re* necessities – necessities that are out there in the world – is a matter of controversy, and raises awkward epistemological questions. Necessities *de dicto*, which are properties of statements, are less controversial, but in the theistic context, the suggestion that 'God exists' is necessarily *de dicto* leads to peculiar difficulties. What I propose to do in this essay is to offer a way of understanding divine necessity which avoids both kinds of difficulty.

In the next section I explain briefly why both of the obvious ways of construing the proposition that God exists of necessity are problematic. In section 3 I outline a projectivist account of necessity, along the lines proposed by Simon Blackburn. Section 4 augments this projectivist account by considering what natural tendencies of thought constrain our imagining of counterfactual situations. Finally, I offer an explanation, in terms of the psychology of ordinary counterfactual thinking, of why the assumption of God's necessity is such a natural one for theists to make. Given even a relatively basic concept of God, one which does not explicitly include a modal element, it somehow goes against the grain of counterfactual thought for the theist to contemplate the proposition that God might not have existed.

2. Two Kinds of Necessity

The first of Quine's three grades of modal involvement consists of the attaching of modal predicates to names of statements (Quine, 1953, pp. 158–9):

$$\text{Nec } `9 > 5'$$

where 'Nec' is taken to stand for 'is necessary' or 'is necessarily true'. For certain statements, their formal structure is sufficient to guarantee their truth, and so they count as necessary: 'If both *p* and *q*, then *p*'; 'If either *p* or *q* and Not-*p*, then *q*'. For another group, the meaning of their terms guarantees it: 'All cubic objects have a shape'; 'A thing is smaller than anything else that wholly contains it.' They are thus true of necessity, but the source of necessity is a feature of language only. This is necessity

de dicto – of the statement. The statements in question are analytically true: their negations would be self-contradictory.[1]

To say that 'God exists' is necessary *de dicto*, therefore, is to say that 'God exists' is necessary simply by virtue of the meaning of 'God'. In other words, some form of the ontological argument is sound. Of course, this is a bold, even reckless, suggestion to make, and most theists would not acknowledge that any form of the ontological argument is sound. But even if there were a sound ontological argument, this would not be good news for theism. For statements that are necessarily true *de dicto* are exceptions to the otherwise plausible thesis that truths are made true by some part of external reality. Analytic statements need no such truth-maker, for their truth is guaranteed by some internal feature of the sentence used to make the statement. So, ironically, if 'God exists' is necessary *de dicto*, then God himself would be *de trop*, not needed for the truth of the statement. To spell this out more fully, if the ontological argument, in some form, is successful, then 'God exists' is analytic. But analytic statements (by themselves) tell us nothing about the world itself. To the extent that they are informative, they tell us only about linguistic conventions. So if the ontological argument succeeded in establishing its conclusion, it would tell us nothing about the world. But, surely, no theist will be prepared to assent to the proposition that the truth of 'God exists' tells us nothing about the world!

If, then, God's existence is necessary, it must involve a different order of necessity. We have to turn to Quine's third grade of modal involvement, in which modal expressions are taken to be sentential operators which are permitted to fall within the scope of the quantifier, so that necessity is represented as a feature of the objects quantified over.[2] For example:

$$(\exists x)\,\mathrm{nec}\,(x > 5)$$

where 'nec' is to be interpreted 'it is necessarily the case that'. The kind of necessity we encounter here is *de re* – that is, of the object. It is a step too far for Quine, who warns that it 'leads us back into the metaphysical jungle of Aristotelian essentialism' (Quine, 1953, p. 176). But, given the self-defeating nature of the ontological argument, if 'God exists' is necessarily true, it must be so by virtue of some feature of God himself, not of the word 'God'.

What feature, exactly? A tempting line of thought is that God's necessity is a consequence of his relation to creation: all things are dependent on him. So if *anything* exists, he does, for he is a precondition of the existence of anything else. Could he nevertheless fail to exist? In such

a world, nothing exists whatsoever: it would be a completely empty world. Now if, as some have argued, it makes no sense to suppose that there could be nothing at all, there is no such possible world. Every possible world, in that case, is a world in which something exists, so every possible world is a world in which God exists.

Tempting though that line of thought is, however, it falls short of offering an account of God's necessity. For it might be thought a contingent matter that everything is dependent on God (much as it might be thought a contingent matter that consciousness requires a physical realization). So we have to build in a modal element to God's relation to his creation: everything else is *necessarily* dependent on God. And if we were puzzled by the idea of a being's having necessary existence, we will be no less puzzled by the idea of creation being necessarily dependent on some other being.

The problem, of course, is not specifically theological: it concerns our understanding of modality in general. Simon Blackburn has suggested that a dilemma faces anyone who takes what he terms a 'truth-conditions' approach to modality, that is, anyone who thinks that the difference between '*p*' and 'Necessarily *p*' is to be explained in terms of a difference in what in reality would make statements of these two forms true. Suppose, on the one hand, we offer an account of what it is for something to be necessary, and we do so in terms that do not involve any modal concepts (i.e. in terms which make no explicit or implicit reference to the necessary or the possible). Then the worry is that we have explained away the modal, and there remains no sense in which the state of affairs in question *has* to be so. On the other hand, we might insist that any analysis of the modal must include the modal: that 'Necessarily *p*' is true if and only if *p* is true in all *possible* worlds, or that '*p*' is contained within any complete and *consistent* description of the world, and so on. Then, although we may have drawn connections between different modal ideas in an informative way, we have, by staying within the sphere of the modal, simply postponed the question of what is conveyed by modalizing (Blackburn, 1987, pp. 53–4).

This dilemma, in turn, is not unique to modality. One can equally raise it about the moral: what is it to say that ϕ-ing is *good*? Or the universal: what is to say that *x* and *y* have the *same* properties? We want an informative analysis, not a circular one. But a truly informative analysis risks getting rid of the analysandum. A typical response to this kind of dilemma is to say that some things are fundamental, and one's philosophical outlook is in large part determined by what one considers to be fundamental. Defenders of *de re* necessity may insist that no further

analysis of the modal is possible, because such necessities are ontological bedrock.

So perhaps the difficulty is not so much that of providing a non-circular analysis of the modal, but rather an epistemological one: just how do we *tell* whether some state of affairs is necessary or not? We do not have to be in the grip of a verificationist account of meaning to feel the pull of the demand that our account of necessity be epistemically responsible. We want to be able to say what difference necessity makes. In the case of physical necessities (elemental sodium reacts with water to form the hydroxide; a magnetized iron bar loses its magnetism when heated etc.) we may not, as Hume pointed out, directly observe any necessity. But we can make an inference to the best explanation: there are these regularities because there is an underlying necessity linking the properties involved. When it comes to metaphysical necessity, however, this kind of inference is not obviously available. What observations could we make, for instance, which would justify the idea that among your *essential* properties (as opposed to your properties, full stop) is the property of belonging to the human race?

Perhaps the demand for empirical confirmation is misplaced, however, and we should look instead to largely *a priori* considerations. As we noted above, modal forms of the cosmological argument provide a reason for thinking of God as necessary (if he exists at all), and the principles on which such arguments are based look more *a priori* than empirical: there is an antecedent cause for every contingent state of affairs; and there is no actual infinite in nature. The second of these principles we might defend by appeal to parsimony and the availability of a mind-dependence account of the infinite. But what is the basis of the first? Is it that whatever allows us to discern contingency in nature also intimates the need for a cause? Or is 'contingent', in this context, just shorthand for 'requires a cause'? Either way, the problem arises how we actually discern contingency (and necessity) in nature.

So the defender of divine necessity faces a dilemma: either the necessity is interpreted *de dicto*, in which case we have a self-defeating ontological argument, or it is interpreted *de re*, in which case we face awkward questions concerning the epistemology of metaphysical necessities. Is there a way of avoiding both horns of this dilemma?

3. A Projectivist Account of Modality

I think there is. The *de re*/*de dicto* distinction is a distinction between objects of necessity, the items to which necessity is ascribed. And it is a

natural assumption to make that this distinction coincides with a distinction between the *sources* of necessity. For when we ascribe necessity to the sentence only, the source of necessary truth is linguistic. And when we ascribe necessity to the world, the source of the necessity is supposed to be some feature of the world. But whereas the distinction between the objects (language and world) is exhaustive, the distinction between these two sources, arguably, is not. A third possible source is the mind. How so? Talking of the necessity that appears to accompany the relation of cause and effect, Hume remarks,

> The efficacy or energy of causes is neither plac'd in the causes themselves, nor in the deity, nor in the concurrence of these principles; but belongs entirely to the soul, which considers the union of two or more objects in all past instance. 'Tis here that the real power of causes is plac'd, along with their connexion and necessity. (Hume, 1739–40, Book I, Part III, Section XIV, p. 166)

The constant conjunction we observe between causes and effects (objects always falling when dropped, changes of speed or direction of motion after collisions etc.) produces, explains Hume, a 'determination of the mind' to pass from the idea of one to the idea of the other – a phenomenon that would later be given the name of classical conditioning. And this determination of the mind, this inevitable transition from the idea of the cause to the idea of the anticipated effect, is then attributed to the external events themselves, so that it appears that there is some necessary connection between the events. Hume thus offers a 'projectivist' account of causal necessity: some feature of the mind is projected, unconsciously, onto the objects themselves. The fact that this projection is entirely automatic and unconscious explains why it should appear to us that we perceive some necessity in the world. If we accept Hume's account, are we obliged to say that our attribution of necessity to causation is illusory? It is perhaps not possible to draw a very sharp contrast between perception and projection, given that much perception is coloured by previous experience and conceptualizing, so there is no need to say that our ordinary causal judgements rest on a mistake. But if anyone insists that causal necessity lies wholly in external relations, then Hume is telling them that they are wrong.

Blackburn extends this projectivist account of causal necessity to other kinds of necessity, even logical necessity.[3] The suggestion is this: when we contemplate, say, arithmetical truths, such as $2+2=4$, we can make nothing of the proposition that they are false, that in fact $2+2=5$.

And this is not simply a failure of imagination. We cannot, for example, imagine what a colour beyond the familiar spectrum would look like, but we would not on that account suppose it impossible that one could have an experience of such a colour. What prevents us from imagining it are the limitations of our actual experience, limitations which are themselves a result of the structure of our visual sense organs. That we cannot see beyond the actually visible spectrum we therefore pronounce contingent. What inclines us to judgements of necessity, in contrast, is the absence of a naturalistic explanation for our imaginative failure. And no naturalistic explanation is in the offing for our failure to make anything of the proposition that $2 + 2 = 5$. It is then that our imaginative block allows us to project the necessity onto the proposition itself (Blackburn, 1987, pp. 69–70). When we turn to geometry, where we do have a naturalistic explanation of our failure to visualize space as having more than three dimensions, or as being curved, we can make something of the idea that physical geometry is a contingent matter.

Moral judgements are an interesting case, standing as they do in an intermediate position between arithmetical judgements and ones reporting physical states of affairs. Once we accept that moral judgements are inextricably bound up with our own emotional reaction to situations, and not simply a cool reading-off of the moral properties exhibited by those situations, we can contemplate the idea of our being differently constituted, having different emotional responses to situations, and so arriving at different moral judgements. But this does not entirely remove our resistance to the idea that, for instance, murder is morally neutral, or positively praiseworthy. The imaginative block is projected as moral necessity – this murder *must* be wrong – but the necessity in question seems less inexorable than arithmetical necessity.

Metaphysical necessities, too, might perhaps occupy an intermediate position between the logically necessary and the physically contingent. The conviction that being human is one of your essential properties could well be a projection of my failure to imagine you as anything other than human – a giant beetle, for instance. But I can also be aware that this failure on my part is tied up with the ways in which we ordinarily identify people, and what is most important in our relations with them. Imagining a giant beetle just gets in the way of thinking about you.

Consider another putative example of metaphysical necessity: the impossibility of backwards causation. Causes never succeed their effects in time. They can only precede or (perhaps) occur at the same time as their effects. We try to imagine causes of backwards causation, of the lighting of a match as being the cause of an earlier explosion, but

find we cannot. The later event seems powerless to bring about what has already happened. The past was what it was, whatever we do now. What might explain this imaginative failure on our part? Here we might plausibly appeal to the connection between perceiving temporal order and acquiring the concept of causation. The fact that our perception of an event is coloured by the memory of an immediately preceding event is what makes us inclined to explain the former by the latter. Once we see this, we might be willing to contemplate the notion of backwards causation, but we will continue to wonder whether we are really making sense of it, or just entertaining a form of words: 'the cause came after the effect'. The necessity is therefore not of the same order as logical necessity, but there is still a difference between the case of backwards causation and the case of non-Euclidean geometry, for instance.

So far, it looks as if the projectivist is providing a causal explanation of the urge to modalize, and nothing more. But it is a short step from this to a more explicit statement of the meaning of modal statements. Recall what Blackburn calls the 'truth-conditions' approach to modalizing: an account of the content of modal assertions in terms of the conditions which make them true – conditions which go beyond those that make non-modal assertions true. The projectivist is able to propose an alternative, 'non-cognitivist' account, according to which modalized statements are not truth-evaluable, since they are purely expressive rather than assertoric: they express a certain attitude towards a proposition, a commitment to it. The difference between 'necessarily *p*' and '*p*' is one of force, rather than truth-conditional content. Blackburn himself resists this non-cognitivist development of the projectivist position, anxious as he is to avoid the famous 'Frege–Geach' objection to expressivism (the objection is directed at moral expressivism, but it is easily adapted to cover other modalities): if we understand 'ϕ-ing is wrong' as merely expressive of an attitude, rather than an assertion, we invalidate arguments in which that expression occurs as a component. The solution is to adopt a 'quasi-realist' view, on which expressions of the problematic kind (moral or modal) are treated as capable of truth and falsity, but their function is still to express a certain state of mind, rather than to draw attention to some extra bit of reality (Blackburn, 1984, pp. 189–96).

That is just the bare bones of a projectivist account of necessity. And it is not yet apparent how it is supposed to apply to the notion of a necessary being, or whether it can apply to such a notion. More needs to be said. Since the account makes appeal to naturalistic explanations

of our imaginative limitations, one way in which it can be augmented is by considering psychological studies of the way in which we ordinarily construct counterfactuals – which non-actual situations we are prepared to contemplate as *what might have happened*, which not, and why.

4. The Psychology of Counterfactual Thinking

Thinking of what might have been is something that apparently comes naturally to us. It is not simply an indulgence, but an important part of learning a skill, cultivating our interpersonal relations, improving historical understanding, avoiding economic mismanagement, and so on. We learn from our mistakes, and the mistakes of others. We also learn from our and others' successes. (What if he had not moved to France that year? Would his business have prospered to the extent it did?) We contemplate counterfactuals quite routinely, and when we do so, we change – in imagination – certain aspects of the past (suppose I had left the house ten minutes earlier), but keep other aspects constant (I walked to the train station). So what governs our choice of what to change and what to keep constant?

Naturally, we are reluctant to contemplate counterfactuals whose antecedents are decidedly *outré* (if I had always lived on the moon, could have transmitted my thoughts directly without using any communicative medium, had been a cactus . . .), but there are apparently some quite ordinary-looking antecedents that are less likely to appear in our counterfactual thought than others. We show certain kinds of bias.

One such bias is temporal. When thinking about a series of independent events that jointly led to a certain outcome, and how those events might have been different, we tend to alter the more recent events in the series. In one study, subjects were presented with the following situation: John and Michael are playing a rudimentary card game for a television show in which they take turns to choose cards which are either red or black. If both pick the same colour card they each win £1000. If they pick different colours, they win nothing. John starts, and picks a red card. Michael (presumably in ignorance of John's choice) then picks a black card, so neither wins anything. Subjects were then asked how things could have turned out differently, and which of John and Michael will feel more guilt. They revealed a tendency to think 'If only Michael had picked a red card, they would have won the cash prize', and that Michael would feel more guilt. In other words, they were more ready to contemplate the counterfactual where the later event was changed, rather than the counterfactual where the earlier event was

changed, that is, 'If only John had picked black, they would have won' (Byrne et al., 2000). By varying the conditions, however, it is possible to compensate for this natural bias. In one variant of the game, John starts and picks a black card. But before Michael picks his, the game is interrupted. After a pause, the game is started from scratch, and John goes first once again, this time picking a red card. Michael follows, picking a black card. Now when subjects are asked how things could have turned out better for the participants, they are just as likely to think 'If only John had picked black' as 'If only Michael had picked red', since the first of these possibilities was made salient by the interrupted game.

In another study, subjects were asked to read a story about a man who decides to stop for a beer on the way home from work, and in addition is delayed by a series of traffic incidents (a lorry manoeuvring, a flock of sheep crossing, a tree trunk blocking the road). He arrives home too late to help his wife who has had a heart attack. On being invited to say how things might have turned out differently in this story, subjects tended to focus on the possibility of his not stopping for a beer, rather than his not being held up in the car (Girotto et al., 1991). Why is stopping for a beer the most natural event to imagine differently? It cannot simply be the time bias operating again, because subjects were given different versions of the story, in which the events occurred in different order. In some versions, stopping for a beer occurred early in the narrative, yet that remained the event most likely to be changed in subjects' counterfactual thinking. The plausible explanation is that that event, unlike the others, was readily controllable. So here we have evidence for another bias in counterfactual thinking: a tendency to adjust events within our control, rather than those outside it.

There may also be a moral dimension. Other studies suggest a bias towards maintaining social or moral norms. That is, people are more likely to contemplate alternatives to actual events where the actual event was morally or socially unacceptable, than ones where the counterfactual alternative involves the unacceptable (Walsh and Byrne, 2005, pp. 71–2). In the narrative above, stopping for a beer broke the taboo against drink-driving, so it is natural for subjects to imagine not having done so.

It is likely that some more fundamental constraint underlies these biases (tendency to change earlier events, controllable events, and those that violate social norms). Our counterfactual thinking tends to be relatively conservative: that is, we make relatively modest changes to the past in contemplating a counterfactual alternative. The reason is not hard to find: the more divergent the counterfactual situation from the

one that actually obtained, the harder it is to work out the consequences. (Just how exactly would your life have turned out if you had been born in the seventeenth century?) In general, in constructing the counterfactual situation, background conditions are kept as they in fact were, and more local adjustments are made to the course of events. The more revisionary the antecedent, the more elusive the consequent – and the greater the demands on our information-processing capacities. That natural conservativeness would explain the tendency to adjust later, rather than earlier events. The earlier in time the adjustment, the more revisionary we have to be, since adjustments have knock-on effects. That, at any rate, is the principle, and although it may not apply in every case, the disposition of mind that it captures will generalize. In the original coloured-card case above, for example, it is actually no more complicated to imagine John's picking black than to imagine Michael's picking red. But the overall habit of the mind will still favour the latter over the former.

Similarly, it is easier to imagine acting differently than it is to imagine events beyond our control turning out differently, because we have a clearer idea, in the former case, of what to change and what to keep constant. In the case of morally salient circumstances, however, there is a check on what actions we imagine differently, for contemplating actions which break taboos involves greater revision than those that do not. A willingness to ignore a particular injunction (e.g. not to jump a red light) raises the question of what other injunctions one might be willing to break. Once we discard certain conventions, or choose to ignore them, much more is up for grabs. In contrast, when we imagine not breaking the injunction that we did in fact break on that particular occasion, we return things to normal, so to speak. What we imagine in that case is more consistent with how we would ordinarily expect/hope to behave.

What significance might any of this have for thoughts about the nature of divine existence?

5. Tracing the Source of Divine Necessity

Consider a relatively basic concept of God, before it becomes subject to demands for modal sophistication: God is the creator of all things, and (in some way yet to be articulated) the basis of all goodness. Let us say that we accept the actual existence of such a being. Is there anything in the way of our contemplating the thought that such a being might not have existed? Let us say that we do contemplate that possibility,

and then make it the antecedent of a counterfactual: 'If God had not existed . . .' . We might treat this as a historical counterfactual, involving some adjustment to the past, and then articulating how what followed would have been different (much as we might construct a counterfactual with the antecedent 'Had the Treaty of Versailles contained no reparation clauses . . .'). In assessing historical counterfactuals we suppose things to be pretty much as they actually were, up until the time at which the antecedent is supposed to obtain. The further back in time we go, the more we have to adjust, but there is at least some background, common to our actual situation, against which the counterfactual situation is cast. But when it comes to contemplating the non-existence of God, this cannot be against *any* determinate background. For the creator is there for the whole of history. So there is no state of affairs prior to God's existence, against which we can cast the antecedent of the counterfactual 'If God did not exist . . .' . In attempting to define the background conditions against which we are to consider God's not existing, we therefore already have to take a view on what the *consequent* of such a counterfactual would be. Would things have come into existence spontaneously? What would their character have been? As theists, we might be at a loss for an answer to such questions. Perhaps nothing at all would have existed. Or perhaps what did exist would be entirely chaotic. Or perhaps it would be as ordered, but as devoid of fundamental purpose, as a machine whose parts are in constant and regular motion, yet which produces nothing. So the proposition that God might not have existed is, for the committed theist, curiously empty: it suggests no very definite situation that we could examine. That does not mean that theists are incapable of contemplating it, but rather that they can make little of it, precisely because so much is bound up with God's existence.

It does not help much to point out that we are, after all, not dealing with an historical counterfactual here. That, since God is outside time, the antecedent 'If God had not existed . . .' is not inviting us to imagine a different past, but the absence of some timeless fact (compare 'If the Inverse Square Law of gravitational attraction had been false . . .'). Again, we have no guide as to what the background conditions for the antecedent are, how much we can, or cannot, take for granted. To imagine how things would have been different in the absence of the creator is akin to asking how *Oliver Twist* would have turned out if Dickens had not written it.

Our reflections in the previous section suggested that counterfactual thinking is naturally conservative when it comes to the amount of revision to the facts in which we are prepared to engage. We are similarly

conservative when it comes to considerations of value: we prefer not to contemplate too many violations of the moral order. But for any theist who views God as the source of goodness, no violation of the moral order could match that of the absence of God, for there would then *be* no moral order. Contemplating the non-existence of God therefore would put the theist at sea, in terms of any value system.

These considerations mirror the (modal) cosmological and moral arguments for God. However, the purpose of those arguments is to establish the existence of God starting from apparently neutral premises about, respectively, the contingency of the universe and the objectivity of moral values. The projectivist account of divine necessity is not an attempt to resurrect those arguments. Rather it points out that, once we adopt the theistic view, the links between God, the universe and value put obstacles in the way of imagining God's possible non-existence.

Of course, for the atheist, things are quite otherwise. Contemplating the non-existence of God creates no intellectual tension since the atheist thinks this is how things are anyway. It is an interesting question, however, whether the quite general constraints on counterfactual thinking we have noted would make it difficult for the atheist to contemplate, or take seriously, the logical possibility of there being a God. Perhaps this too would involve too great a violation of both fact and value for the atheist to give it more than a rather abstract content.[4]

I have suggested that a natural predisposition to counterfactual conservatism makes the thought of God's non-existence less readily available to the theist than thoughts about other kinds of counterfactual situation. This difficulty in making much of the thought of God's non-existence might then be projected onto God himself, giving rise to a further thought that is expressed by the proposition that God's existence is necessary. In saying that God is necessary, we need then be adding no further factual content to the non-modal assertion that God exists; the difference, rather, is one of force. But to this line of thinking, two objections could be raised. The first is that people's thinking about God often shifts, from theism to atheism, or vice versa. Take, for example, an atheist who, as a result of either religious experience, or careful consideration of theistic arguments, comes to believe in God. Prior to conversion, he or she had no difficulty in contemplating the possibility, indeed the actuality, of God's non-existence. Is it plausible that, now a believer, this thought should no longer be available? The second objection is that theists routinely defend their position by pointing to the unattractive consequences of the opposing hypothesis, and this suggests that they are able to entertain and assess the associated counterfactual

'If God did not exist . . .'. They must therefore be able to make something of the possibility of God's non-existence.

On the first of these objections, we might note that thoughts we entertain at some times can indeed become less accessible at others, as a result of a shift in our outlook. We may well remember that we used to think such-and-such, but be unable to capture the sense of its obviousness or inevitability, and perhaps even struggle to make sense of it. Consider shifts in moral thinking, for example. Suppose Helen, at the age of 20, regards euthanasia as murder, pure and simple. However well motivated it might be on some occasions, she cannot reconcile herself to the idea that it is at all morally acceptable. But thirty years later, after a career in health care, looking after the terminally ill, she sees it in quite a different light. She knows that she once thought it wrong, but the conception she now has of what it is for something to be right or wrong makes it a struggle to conceive of how those acts of euthanasia that are performed out of love and compassion for the suffering patient could possibly be wrong. Similarly, the convert may view his or her earlier conviction of God's non-existence as not merely wrong, but muddled.

On the second objection, to express the importance of a certain truth in terms of the conditional, which says what would be the case if it were not true, is not necessarily to concede the intelligibility of the antecedent. Consider this conditional: 'If there were no space, nothing would have mass.' That is one way of drawing attention to the connection between the property of having mass and the property of occupying space. The thought conveyed is perfectly intelligible. It does not follow that a world without space is intelligible. Similarly, one might say, 'If there were no irrational numbers, then Pythagoras's Theorem would be false', meaning simply that Pythagoras's Theorem yields irrational values for the hypotenuse, given certain values for the other two sides. It implies neither that the absence of irrational numbers is a genuine possibility, nor that we can clearly understand what it would be for there to be no irrational numbers. What I want to suggest, then, is that there is a difference between expressing certain thoughts by means of counterfactuals, and engaging in genuine counterfactual thinking. The latter requires, as the former does not, an understanding of the situation represented by the antecedent as a fully intelligible state of affairs, and working out what would be the case in a world in which that state of affairs obtained. Similarly, a theist may express the deep connection they see between God and the moral order by means of the counterfactual 'If God had not existed, there would have been no moral order', without thereby implying that the antecedent represents a

situation they can make perfect sense of. Counterfactual assertion does not always betoken genuine counterfactual thinking.

6. Conclusion

In summary, then: Why is it so natural for the theist to say that God exists of necessity? For the theist, the domains of fact and value are pervaded by God: nothing is fully intelligible without reference, at some point, to God. This, together with a natural conservatism in counterfactual thinking, a conservatism revealed by psychological experiments which require subjects to imagine how things might have turned out differently in certain contexts, makes it harder, more unnatural, for the theist to imagine seriously a possible world in which God does not exist. And this difficulty experienced by the theist in making much of the notion of God's non-existence is then projected onto God, in the form of the idea that God cannot fail to exist. As we noted earlier, however, not all imaginative blocks are projected onto the world. Our inability to imagine a fourth primary colour is not projected as the impossibility of there being more than three primary colours, since there is a naturalistic explanation of why we see colours as we do. But have we not just given a naturalistic explanation of the theist's difficulties in imagining a world without God? Well, we have, but there is still a difference between this kind of explanation and the explanation of our imaginative limitations over colour. Although we cannot imagine a fourth primary colour, we can readily conceive of beings with differently constituted visual systems. It is not so straightforward for the theist to imagine a being who is in a better position than he or she is to conceive of a world without God. Belief in God is not simply accepting the existence of another item in the landscape: it completely transforms that landscape. And that transformed landscape presents itself as a vision of how things are, not simply of how they appear from a certain perspective. The difficulty of conceiving of God's non-existence appears more objective than the difficulty of conceiving of a fourth primary colour, as does conceiving of alternative arithmetics. From the committed theist's point of view, any rational thinker, however physiologically or neurologically constituted, will struggle genuinely to understand the absence of God.[5]

That provides a causal explanation of theological modalizing, but it also offers something more: an account of what it is to ascribe necessity to God. The centrality of God to the theist's conceptual framework, its role in determining the character of that framework, gives rise to a commitment that seems unavoidable. 'God necessarily exists' expresses that

commitment. And, to borrow from Blackburn's account of modal thinking, we do not have to deny that the expression has assertoric force, that it is evaluable as true or false. The sense of the inexorability of commitment is what provides assertibility conditions for the statement.

Before leaving the topic, there is a final, methodological worry to be addressed. I have appealed to natural biases in our construction of counterfactuals to explain the psychological basis of the doctrine of divine necessity. But, it will be urged, the philosophical analysis of divine necessity surely transcends such natural dispositions. So whereas psychology might cast some light on the aetiology of theological doctrine, it can hardly illuminate the true meaning of that doctrine. In reply, I would point out that philosophy (in this context) is in the service of theology, not vice versa. And theology, in turn, is an attempt to articulate our deep-seated beliefs about the nature of God. Where those beliefs come from is not irrelevant to understanding their meaning. Faith first, theory second.

What has been offered here is a projectivist account of God's necessity. It is not a projectivist theory of God. We have assumed, that is, that there is a being onto whom necessity is projected, just as, according to Hume, causal necessity is projected onto events in the world. The natural question this prompts, namely what other divine properties might also be projected, I leave for the reader to ponder.

Notes

1. Quine himself does not characterize the first grade of modal involvement in terms of analyticity, having elsewhere cast doubt on the analytic/synthetic distinction (Quine, 1951).
2. The second grade of modal involvement, which we have skipped over here, takes modal terms as operators attaching to statements:

   $$\mathrm{nec}\,(9 > 5)$$

   The inverted commas from the first grade have disappeared, implying that it is the content of the statement that is modified by the modal expression. But we do not yet have quantification outside the scope of the operator. As Quine explains, the second grade can be understood as a way of articulating the relatively harmless first grade, but it can also be a stepping stone to the anything but harmless third grade (Quine, 1953, p. 176).
3. Why do we need a projectivist account of logical necessity? Is it not enough to appeal to logical laws and linguistic conventions? Quinean scruples, however, may make us wary of drawing a sharp *de dicto*/*de re* distinction, insofar as the former is taken to be coincident with the analytic. And in the case of arithmetical truths, we will not be able to combine a Platonist ontology (i.e. one that holds that there really are numbers, independently of arithmetical

thought and language) with the view that arithmetical necessity does not go beyond the level of language.

4. This might seem to raise the spectre of incommensurability. If what theists assert does not coincide with what atheists deny, the prospects for rational debate in this area are dim. However, the inability of each side fully to enter imaginatively into the other side's point of view does not mean that they cannot articulate the propositions over which they disagree.
5. As Anselm put it, 'How indeed has he [the Fool] "said in his heart" what he could not think . . . in one sense a thing is thought when the word signifying it is thought; in another sense when the very object which the thing is is understood. In the first sense, then, God can be thought not to exist, but not at all in the second sense' (*Proslogion*, Chapter IV; Charlesworth, 1965, pp. 119–21).

References

Blackburn, Simon (1984) *Spreading the Word: Groundings in the Philosophy of Language* (Oxford: Clarendon Press).

Blackburn, Simon (1987) 'Morals and Modals', reprinted in Blackburn, *Essays in Quasi-Realism* (Oxford: Oxford University Press, 1993), pp. 52–74.

Byrne, R. M. J., S. Segura, R. Culhane, A. Tasso and P. Berrocal (2000) 'The Temporality Effect in Counterfactual Thinking about What Might Have Been', *Memory and Cognition* 28: 264–81.

Charlesworth, M. J. (1965) *Anselm's Proslogion* (Oxford: Clarendon Press).

Findlay, J. N. (1955) 'Can God's Existence Be Disproved?', in Antony Flew and Alasdair MacIntyre (eds), *New Essays in Philosophical Theology* (London: SCM Press, 1955), pp. 47–56.

Girotto, V., P. Legrenzi and A. Rizzo (1991) 'Event Controllability in Counterfactual Thinking', *Acta Psychologica* 78: 111–33.

Hume, David (1739–40) *A Treatise of Human Nature*, ed. L. A. Selby-Bigge, rev. P. H. Nidditch (Oxford: Clarendon Press, 1978).

Quine, W. V. (1951) 'Two Dogmas of Empiricism', *Philosophical Review* 60: 20–43.

Quine, W. V. (1953) 'Three Grades of Modal Involvement', in idem, *Ways of Paradox* (New York: Random House, 1966), pp. 158–76.

Walsh, Clare R., and Ruth M. J. Byrne (2005) 'The Mental Representation of What Might Have Been', in David R. Mandel, Denis J. Hilton and Patrizia Catellani (eds), *The Psychology of Counterfactual Thinking* (Abingdon: Routledge), pp. 61–73.

2

Why Would Anyone Believe in a Timeless God? Two Types of Theology

Benjamin Murphy

1. Is Theology a Waste of Time?

God may have infinite time, but for the rest of us, time is a precious commodity, to be used carefully. Consequently, it is somewhat disturbing to contemplate the possibility that theology is a complete waste of time, particularly if one is a professional or even an amateur theologian. No doubt practitioners of other disciplines have such feelings on occasion. The difference is that recently the proposition that theology is a waste of time has been advanced as a thesis that has scientific support.

I am referring here to research that has been carried out by practitioners of that relatively new discipline, cognitive science of religion. The overwhelming view within this field is that religion itself is a natural human phenomenon. According to this view, a tendency to believe in the supernatural is part and parcel of being an average human being. It is not that religious belief is innate, but that, from an early age, children are primed to accept religious beliefs. If cognitive scientists are to be believed, religion has a healthy future.[1]

However, along with good news for religion comes bad news for theology. It appears that there is frequently a gap between the official beliefs propounded by religious professionals, such as theologians, and the beliefs that come naturally to most people. Furthermore, in day-to-day thinking, even those who have had theological training tend to revert to their natural religious beliefs, a phenomenon known as 'theological incorrectness'. In one study (Barrett and Keil, 1996) participants were given a series of stories about God and asked to retell them. When retelling the stories, people have a notable tendency to describe God

in more anthropomorphic terms. For example, theologians agree that God can perceive everything, but a story in which God is listening to a bird while an aeroplane lands became a story in which the sound of the aeroplane landing interrupts God's enjoyment of the birdsong. God, if he exists, is surely capable of multi-tasking, but a story in which God answers listens to and answers two prayers becomes a story in which God deals with one prayer at a time. When God is replaced by a specially created fictional entity with super-abilities that include super-hearing or multi-tasking, people do not make these errors.

These are perhaps trivial errors in a trivial task. Ever since God was described as walking through the Garden of Eden in the cool of the day, storytellers have been willing to spice up stories with a bit of anthropomorphism. The anthropomorphisms that people resort to in the experiment can be said to improve the story, since they give a rationale for otherwise irrelevant details: if the noise of the aeroplane wasn't an interruption, why even mention it? As Chekhov pointed out, if you show the audience a gun in Act One, it must be fired by the end of the play. In other cases, theological incorrectness is not so trivial. For example Theravada Buddhists may say that the Buddha was a normal human being who attained a goal, enlightenment, that is open to any human to achieve, but their refusal to ordain nuns now that the valid lines of female ordination have been broken is best explained by the lingering belief that there was something supernatural about the touch of the Buddha's hands (Slone, 2004, pp. 81–3). The theologically incorrect belief about the Buddha's supernatural touch motivates the politically incorrect refusal to ordain women.

So, it seems that religious beliefs occur naturally without any help from theologians. When theologians try to introduce ideas that conflict with these natural religious beliefs, the best they can hope for is lip-service. So, why do we keep on trying? This is not only a source of dismay for theologians, it is a puzzle for cognitive scientists. Why would so many intelligent people persist, over many generations, in an activity that is doomed to fail? Why spend so much time trying to teach difficult theological systems when there is the easier option of indoctrinating people with naturally attractive religious beliefs?

Of course, to say that theology is a failure because theologians do not manage to win most people over to their beliefs presupposes that the main goal of theologians is to attract as many adherents as possible. Impact and popularity are factors that can be objectively measured, and it is understandable that cognitive scientists of religion adopt such a standard to measure the relative success or failure of religious writers.

Yet this misses the fact that theologians, like many academics, may not be motivated by a desire to win popular support. Rather than judging theology a failure by standards imposed from outside, we should attempt to understand what factors, other than the desire to persuade large numbers of people, might motivate people to become theologians.

Joseph Bulbulia suggests that it is the fact that the study of theology requires a lot of time and yet has no value outside a religious community which makes it useful as a way of signalling a high level of commitment to the community, and speculates that it may serve 'few utilities' apart from this (Bulbulia, 2009, p. 52).

Bulbulia writes this because he believes in what Pascal Boyer refers to as the 'tragedy of the theologian'. According to Boyer (2001, pp. 283–5), religious professionals frequently organize a 'guild of priests' to protect their interests. The guild claims a monopoly on knowledge of the supernatural, and the agreement of all members of the guild on the contents of such knowledge is evidence of their collective reliability. That level of stability requires a system, thus theology comes into existence, but teaching people systems requires repetition, and this induces tedium. People are inevitably drawn to the new, the exciting, that is to say, to the unorthodox. Theological orthodoxy will never prevail.

As Boyer points out, the distinction that he is employing between routine, bureaucratized religion and charismatic, revolutionary religion is not new; it can be traced back to the work of Max Weber. However, Boyer is relying on a recent proposal by the anthropologist Harvey Whitehouse that is intended to provide a firm scientific basis for this familiar distinction (2001, p. 283). Whitehouse points out that any long-lasting religious idea must be implanted somehow in our memory, and, given what we now know about the way human memory functions, this can happen in one of two ways. Either a formula is repeated again and again, like learning multiplication tables or conjugating verbs, or else a lesson is associated with some vivid, unusual and exciting event. Whitehouse calls the former technique of teaching and learning a religion the doctrinal mode of religiosity. It is this mode of religiosity that is utilized by the guild of priests, according to the theory, and tedium is an inevitable side-effect. The latter technique of learning, that of creating exciting events, is associated with Weber's charismatic religion, or, as Whitehouse calls it, the imagistic mode of religiosity (Whitehouse, 2004, pp. 63–85).

Boyer and Bulbulia both write as though there were nothing more to theology than the composing and memorizing of catechisms, as though good theology is nothing more or less than orthodox theology. Once again, this is an outsider's view of theology. A five year old reciting

the four-times table may well imagine that, at university level, maths students recite the seventeen-times table. In the same way, Boyer and Bulbulia seem to picture theology as being simply a very advanced version of the kind of Sunday-school classes that were satirized by Mark Twain. This leaves out the creativity that characterizes theology at its best – a creativity that is quite independent of orthodoxy. For example, I have heard conservative Catholics admit that Hans Küng and Karl Rahner excelled in the field of theology, while regretting their lack of orthodoxy. Indeed, one sometimes encounters the idea that excessive zeal for theology tends to lead to heresy. One might try to defend Boyer's thesis by classifying such renegade theologians with Weber's charismatic prophets, breaking away from the monotony of doctrinal religion. However, this would not be in keeping with the scientific basis for the distinction between two modes of religiosity that Boyer takes from Whitehouse. What is distinctive about the imagistic mode, which corresponds to Weber's charismatic religion, is that memories are formed by providing, on rare occasions, striking rituals that appeal to the senses. Certainly, one could describe Karl Barth, for example, as a prophetic voice and a charismatic writer, whose work is full of verbal fireworks. Only the use of literal fireworks, however, would qualify Barth as a practitioner of the imagistic mode of religion in Whitehouse's sense. I suggest, then, that theology is a creative discipline with internal standards of excellence that are independent of the standards by which orthodoxy is judged. I will not argue this point at length, since I trust that readers of this volume are capable of judging the truth of this for themselves.

This means that, from the perspective of cognitive science of religion, the practice of theology remains an unsolved puzzle. There is one other suggestion about the role of theology that comes not from a cognitive scientist, but from a philosopher with a strong interest in cognitive science of religion, Daniel Dennett. He states,

> A milder and more constructive response to relentless skepticism is the vigorous academic industry of theological discussion and research, very respectfully inquiring into the possible interpretations of the various creeds. This earnest intellectual exercise scratches the skeptical itch of those few people who are uncomfortable with the creeds they were taught as children, and is ignored by everybody else. (Dennett, 2006, p. 208)

Dennett is not in the business of flattering the egos of theologians, but he does at least acknowledge that theologians have genuine intellectual

motivations, beyond a desire for status within or control of a community. There is, after all, a certain kinship between philosophers and theologians. It is not unknown for philosophers to advance abstruse theories that defy common sense, and not usual for them to retract such theories when they fail to attract a popular following. I think that Dennett's description does fit much analytical philosophy of religion, and analytical theology. However, as I hope to demonstrate, it does not cover the whole of theology.

According to cognitive science of religion, popular religious concepts stick in our mind because they are minimally counter-intuitive. The term 'minimally counter-intuitive' was introduced by Justin Barrett, based on the work of Pascal Boyer (Barrett, 2004, p. 30). A concept is minimally counter-intuitive if it violates a single ontological category: flying fish, flightless birds, virgin births and men who walk on water are all examples. These concepts stick in the mind because they violate an ontological category, and thus contain an element of surprise. However, because they violate only one ontological category, we can still form expectations based on our existing knowledge. A flightless bird still has feathers and a beak, and a man who walks on water still has two legs (Boyer, 2001, pp. 62–5). Theological concepts, unlike popular religious concepts, violate many expectations and are thus harder to grasp. The question is why anyone would even try.

In this essay I have picked one particular property that separates the abstract and extremely counter-intuitive God of theology from the anthropomorphic and minimally counter-intuitive God of popular religious belief: divine timelessness. I plan to consider other divine attributes, such as impassibility, in future work. If a God who multi-tasks is hard to understand, how much more so a God who is completely outside of time? What is more, even advocates of divine timelessness, such as Paul Helm, admit that fidelity to the biblical tradition does not require such a concept (Helm, 2001, pp. 31–32). In 1970, Nelson Pike concluded his study of *God and Timelessness* with a question: 'What reason is there for thinking that God's timelessness should have a place in a system of Christian theology?' (Pike, 2002, p. 190).

Pike is anything but ignorant of the arguments that have been advanced in favour of divine timelessness. His question comes after 190 pages of careful analysis. So his question is not about the justifications that have been advanced for believing in a timeless God, but the underlying motivations. Frequently, when people construct complicated philosophical arguments, it is because the conclusion is

intuitively attractive. What attraction might the doctrine of divine timelessness have?

2. Timelessness and Foreknowledge: Theology as a Technical Discipline

In the *Summa Theologiae* Iae, Q. 14, Art.11, Aquinas considers the question of God's knowledge of future contingents. To summarize his reasoning: since God knows now what actions humans will perform, there are truths about future choices, but since such future choices are contingent, there cannot now be truths about them. The solution is that God is timeless. Our future is part of God's present. God knows everything, but his knowledge is simultaneous with what is known.

Anthony Kenny has pointed out the following problem. Simultaneity is a transitive relationship. If A is simultaneous with B, and B is simultaneous with C, then A must be simultaneous with C. But in that case, since God's knowledge of every free action performed by humans is simultaneous with those actions, all human actions must be simultaneous with each other (Kenny, 1979, pp. 38–9). A minimally counterintuitive idea is still comprehensible enough to generate a set of expectations and inferences that make sense. However, as Kenny's argument shows, the idea of a timeless God generates inferences that undermine all of our usual chronological expectations. It is counter-intuitive, certainly, but far more than minimally so.

Both Aquinas's initial argument and Kenny's objection follow the same pattern of reasoning, a methodology that can be traced back to Aristotle, and remains common amongst analytical philosophers today. Here is the methodology as explained by Aristotle in the *Nicomachean Ethics*:

> As in the other cases we must set out the appearances, and first of all go through the puzzles. In this way, we must prove the common beliefs about these ways of being affected – ideally all the common beliefs, but if not all, then the most important. For if the objections are solved and the common beliefs are left, it will be an adequate proof. (1145b, 5; Aristotle, 1999, p. 100)

For Aquinas, the common beliefs are that human actions are contingent and that God has infallible knowledge of them. The puzzle is that this creates a contradiction, and then the doctrine of divine

timelessness – a rather uncommon belief – solves the problem. Kenny then brings in another common belief – that simultaneity is a transitive relationship – and creates another problem, and so it goes. It is, as I have said, a method frequently employed by analytical philosophers. To give just one example of this methodology as practiced within analytical philosophy, Donald Davidson reasons that common beliefs tell us that human actions are unpredictable, and also that all events are, in principle, predictable, and his theory of anomalous monism offers a reconciliation (Davidson, 2001). So it is not surprising that Daniel Dennett, an analytical philosopher, should look upon theology as an exercise that follows this same pattern. The clash in the common beliefs creates a sceptical itch, an awkward philosophical question, and the theologian aims to provide an intellectually satisfying answer. Indeed, Ilkka Pyysiäinen has proposed this as a model for how theology is created (Pyysiäinen, 2004).

Pyysiäinen points out that many cognitive theorists have offered a dual-process view of the human mind, according to which we have two ways of thinking, one of which has been labelled 'implicit', 'experiential' and 'intuitive' and the other which has variously been called 'explicit', 'rational' and 'reflective'. Despite the different terminology used by different theorists, it is clear that they are all referring to the same distinction. To avoid unwarranted inferences that might come from adopting any particular labels from his list, Pyysiäinen refers to implicit/experiential/intuitive thinking as the A-system, and explicit/rational/reflective thinking as the B-system (Pyysiäinen, 2004, p. 142). As the reader has no doubt realized, Pyysiäinen's A-system is closely related to Whitehouse's imagistic mode of religion, and his B-system to Whitehouse's doctrinal mode (Pyysiäinen, 2004, pp. 127–8).

It is the A-system that supplies what I have been calling, following Aristotle, common beliefs. The main finding of cognitive science of religion, as I have already noted, is that the A-system will naturally produce religious beliefs. The Aristotelian task of sorting through these beliefs and solving puzzles falls to the B-system.

Where Pyysiäinen differs from Boyer is in his suggestion that cognitive science needs an integrated model. As far as Boyer is concerned, theology, which is doctrinal, is perpetually engaged in a losing battle with imagistic thinking. Pyysiäinen does not see conflict, rather he states that the A-system and B-system 'work together most of the time' (Pyysiäinen, 2004, p. 127) and on the basis of this integrated model, he hopes to tackle the issue that he thinks Boyer avoids: the origin of theological ideas (Pyysiäinen, 2004, p. 143). So, according to Pyysiäinen,

theology is what happens when the B-system attempts to rise to the challenge of the minimally counter-intuitive ideas that the A-system sends its way:

> To the extent that counterintuitiveness is based on explicit modification of intuitive concepts, it is to be expected that theological thinking is characterized by both counterintuitiveness and a tendency to rationalize. Theologians, indeed, try to rework minimally counterintuitive ideas into such abstract concepts that the apparent contradictions seem to disappear. (Pyysiäinen, 2004, p. 143)

If this is what theologians do, then successful theology becomes a matter of carrying out the balancing act described by Aristotle. As many of the common beliefs as possible must be preserved, provided the puzzles can be solved to the satisfaction of awkward philosophers like Kenny. Stump and Kretzmann's solution to Kenny's puzzle is to introduce a highly counter-intuitive (but, they claim, logically consistent) concept of simultaneity which is non-transitive (Stump and Kretzmann, 1981). Open Theism, on the other hand, sacrifices the common view that God has infallible knowledge of all future actions in the interest of maintaining a coherent system. I will refer to theology that follows this pattern as 'technical theology'.

It is important to be clear that technical theology is never purely technical. The technical aspects – the work of the B-system – are meant to support, as far as possible, the non-technical intuitions of the A-system. This may be compared to the work of a software engineer. I use Microsoft Windows on my computer, and as long as it runs without any problems, I am happy. But sometimes problems do arise, and I have either to solve them myself, or get in a software engineer to help. In the old days, fixing problems sometimes required the use of MS-DOS, which was much less user-friendly. I was aware that MS-DOS was there, and I knew how to enter simple instructions, but as long as my computer was functioning smoothly, I was happy to ignore its existence. Stump and Kretzmann's theory of ET-Simultaneity is rather like MS-DOS. Only a tiny minority of believers are even interested in trying to understand it. If those who do understand it conclude that it is a coherent solution to the puzzle, then they can reassure the majority of believers that there is a solution to the problem of divine foreknowledge, the exact details of which are of interest only to the technically minded. If the theological boffins reject Stump and Kretzmann's theory, then they might have to resort to Open Theism, which does interfere somewhat with the natural inclinations of

believers – as though a software engineer were to tell me that I can run Windows on my computer, but only in a foreign language. Either way, the technical stuff is not an end in itself: the B-system is used to organize, support and repair the indispensable raw material provided by the A-system. Or, in more familiar theological terminology, technical theology tends to serve an apologetic function: it defends and occasionally modifies an existing structure of beliefs, but does not create a new one. It will not have escaped attention that what I am calling 'technical theology' is a common approach amongst analytical philosophers of religion. Indeed, if I am right, Pyysiäinen has provided a model that explains the way a lot of philosophical and theological thinking functions.

However, I do not think that Pyysiäinen's model is sufficient to explain all theological thinking. When Aquinas was writing in the thirteenth century theology was established as a rigorous academic discipline. But the doctrine of divine timelessness entered into Christian thinking well before theology became part of the scholastic curriculum, and it was not motivated purely by technical concerns. To understand the real attraction of this idea, we must trace it back to the pre-scholastic era.

3. Timelessness and Modes of Consciousness: Theology as Experience-Enhancing

When Aquinas affirmed that God is timeless, he was reinforcing a tradition that had already been strongly established, Boethius's *Consolations of Philosophy* clearly standing out as a particularly influential source. Boethius, of course, did not invent the idea of a timeless God; it was taken from Plotinus. Nor was he the first Christian theologian to make use of the idea; it had already been integrated into Christian theology by Augustine. But still, as John Marenbon states, the *Consolations* has exerted a magnetic pull on readers because it tackles the theme of human happiness within the most dramatic of contexts (Marenbon, 2003, p. 96). If we want to know why the idea of a timeless God has been so popular for much of the history of Christian theology, Boethius is surely the author to whom we should turn.

When Boethius wrote *The Consolations of Philosophy*, he was in prison awaiting execution on a charge of treason. The drama of the book comes not merely from his impending execution, but the fact that he fell so suddenly from high political office. His unhappiness is increased by the thought of everything he has lost, and at first, he gives way to lamentation. He is interrupted by the arrival of a lady who personifies

Philosophy. She banishes the Muses of poetry, who have encouraged him to wallow in emotion, and reminds him that he took up the study of philosophy because he aspired to true happiness, which he should have known could not be attained from wealth and status that is here today and gone tomorrow. Philosophy then assures him that God is true happiness, and that we humans become happy insofar as we become more like God (Boethius, *Consolatio Philosophiae*, Libro III, Prosa 10). Later, Boethius gives his famous description of what God's life is like. God is eternal, and eternity is 'complete and simultaneous possession of endless life' (Boethius, *Consolatio Philosophiae*, Libro V, Prosa 6).[2] Since God is true happiness, and we become happy insofar as we become like God, it follows, one would expect, that the closer we come to a complete and simultaneous perception of endless life, the happier we will be. Perhaps we could have, if not a complete and simultaneous possession of endless life, a nearly complete and simultaneous possession of (segments of) our own finite life. In that case, Boethius's high status would no longer be perceived as something that once was his but now is lost.

But what would it mean to experience different times as present?

Here I turn to the work of Margaret Donaldson, one of the leading figures of developmental psychology. In her book *Human Minds* (Donaldson, 1992) she traces the development of the modes of consciousness we have available to us as we reach maturity.[3] Our first experiences are in 'point mode', focusing on what is taking place now. To say that things are 'taking place' is dynamic, not static: the 'point' is 'not like the idealized extension-less point of the mathematician but rather like a moving spot of light' (Donaldson, 1992, p. 42). As we grow older, we start to make definite plans about the future, drawing upon our knowledge of the past. This involves thinking in what Donaldson terms the 'line mode', a process that begins at around eight to ten months (Donaldson, 1992, p. 61). Line-mode thinking is important because it enables us to become responsible. If I understand Donaldson's position correctly, line-mode thinking is valuable because it brings to light errors that we might otherwise miss. Even though right now I am safe and comfortable, perhaps my actions of this morning are likely to have bad consequences for me tomorrow, unless I rouse myself to take action now. This explains why Donaldson associates the line-mode with negative emotions. She describes it as 'the mode of "if only" and "what if" – of resentment, guilt and fear' (Donaldson, 1992, p. 61). Like a car alarm, the line-mode alerts us to hidden threats and keeps us awake at night.

The other two modes of consciousness are the construct mode, in which we imagine what might happen in some unspecified time

(Donaldson, 1992, p. 81),and the transcendent mode in which our thoughts are not bound to any place or time (Donaldson, 1992, p. 125). If I imagine a mathematical problem involving two trains setting off from two unspecified stations, that would involve the construct mode. I am thinking of a concrete train travelling on a journey, but not a particular train on a particular journey. If I then focus purely on the mathematics, and forget about the train itself, I enter the transcendent mode. This particular choice of example makes it sound as though the construct and transcendent modes are purely intellectual, but this is not necessarily the case. Donaldson states that the construct and transcendent mode can be value-sensing rather than intellectual (1992, pp. 150–9), where value-sensing involves arriving at judgements about what is worth paying attention to and what may be ignored (pp. 13–14).

Although Donaldson does not specifically mention Boethius, she is well aware that the history of spirituality provides data that support her theory about modes of consciousness. She points out, for example, that the goal of Buddhist meditation seems to be the attempt to escape from the tyranny of the line-mode and return to the point-mode, thus having the sensation of transcending time (1992, pp. 222–3), and she is aware of the connection to the idea of divine eternity (p. 157). Her reflections on the connections between modes of consciousness and religion are deeply indebted to the work of Keith Ward, whom she cites in the following passage:

> Ward says that such a sense is best evoked not by science or philosophy but rather by poetry or music; 'it is a major heresy of post-Enlightenment rationalism to try to turn poetry into pseudo-science, to turn the images of religion, whose function is to evoke eternity, into mundane descriptions of improbable facts'.[4]

There is a danger that Donaldson's gloss will obscure Ward's real point. Her passage might be taken to suggest that music and poetry are interchangeable, but they are not. Pure music, without words, can evoke moods and emotions and even, in its way, engage the intellect, but it has no propositional content. To get propositional content, we need to add lyrics – and if we detach the lyrics from the music, we are left with poetry. Once we have words, we have the possibility of propositional content. Theology is not pseudo-science, but nor is it pseudo-music: theologians should not try to compete with musicians in the realm of pure emotion. Theology, like poetry, does offer propositions, but propositions that need to be understood within a particular emotional context.

In the *Consolations*, we can, to a certain extent, track the movement between imaginative and factual writing, because Boethius moves between prose and poetry. The qualifying phrase, 'to a certain extent', is important. The dividing line between prose and poetry is permeable. After all, the very idea of a conversation with a personification of Philosophy is clearly metaphorical. However, it is surely not a modern heresy to say that some of the propositions that Boethius ascribes to Philosophy really are intended as serious philosophical propositions.

The initial section in which Boethius gives way to sorrow is written in verse. The arrival of Philosophy is marked by a shift to prose, and her first act is to banish the Muses of poetry. From then on, although the book continues to alternate between prose and poetry, it is clear that the prose sections of the book, in which Boethius and Philosophy engage in vigorous intellectual debate, are driving the narrative forward. Poetry and imagination have their place, but they are subordinate to reason, that is Philosophy, who provides the direction. Within the poems, the gods of Greek and Roman mythology make frequent appearances. Since the author is a Christian, we know that in this context stories of Bacchus, Circe and Orpheus are not intended as descriptions of improbable facts. Within the prose sections, pagan deities are almost entirely absent. There is indeed a reference to Jove's two vessels, one containing good and one evil, in Libro II, Prosa 2, but these are referred to as something that Boethius was surely told of when he was a young man, rather than being referred to without qualification as really existing objects. Indeed, I would go so far as to suggest that one function the poetry serves is to emphasize, by means of contrast, the factual nature of the philosophical propositions advanced in the prose sections.

When contrasting 'facts' with metaphors, we often refer to 'hard' facts, or 'brute' facts. Facts are the things we will have to face up to sooner or later whether we want to or not, and by describing them as hard, we indicate that perhaps we may not want to. When Samuel Johnson attempted to refute Berkeley by kicking a stone, he completely misunderstood the point of Berkeley's philosophy, but he did exhibit a grasp of what most people take to be the distinguishing feature of reality: it gets in your way. In Libro II, Prosa 1, when Philosophy reminds Boethius that in the past he used 'manly words', her point is that he should face his troubles with courage, head-on. If her consolation is really to have the value that she claims for it, it must be such as to enable Boethius to better understand his position on the basis of solid facts, rather than merely distracting him with airy persiflage lacking in any real content.

In emphasizing that the prose passages contain factual content I am not trying to deny that there is also emotional content. My suggestion is not that the switch from poetry to prose indicates a turning away from emotion as such, rather, it indicates a shift in emotional register as Boethius's writing adopts a suitably manly tone. The emotions are those of a warrior, boldly staring the enemy down. The prose sections are factual not because they are written from a neutral standpoint, but because Boethius is determined not to flinch in the face of reality.

The application of Donaldson's modes of consciousness to the *Consolations* should, by now, be apparent. Boethius's situation is a bad one, and it is made all the worse by the fact that he is stuck in line-mode, reviewing his past decisions and preparing for the future. Line-mode thinking is useful when we are trying to take responsibility for our lives and make plans, but Boethius has now lost the ability to control his own life. Line-mode thinking only accentuates his frustration. There are no more plans to be made, and so he must turn his mind to something else. Initially, he turns to the value-sensing mode, as he explores the question of what is truly valuable in life. However, his answer, that what has always truly mattered to him is the pursuit of truth by means of philosophy, leads him naturally into the intellectual modes, and by the final book, where he considers the abstract issue of God's foreknowledge, he becomes caught up in the intellectual transcendent mode. The intellectual transcendent mode offers an escape from the line-mode, and thus carries with it a sense that we are stepping outside of time. Boethius's goal, of course, is to attain true happiness by becoming more like God, a being whose state of mind – grasping everything all at once – is precisely that of someone in the intellectual transcendent mode, as all the different pieces of the puzzle are fitted into their proper place. Ward's statement about what it means to have a vision of God can here be applied to Boethius; it is a vision 'of one's self, as transfigured by the infinite' (Ward, 1987/1998, p. 153).

The use of intellectual problems as a means of coping with stressful situations is not, of course, unique to theology. In 1908, the German industrialist Paul Wolfskehl made plans to commit suicide at midnight precisely. With everything prepared, he decided to spend his last remaining hours studying a paper that dealt with Fermat's infamous last theorem. He noticed an error in the paper he was reading, and became engrossed in finding a way to correct it. By dawn, he had completed the correction and made the choice not to commit suicide. Later, he established a prize for anyone who could definitively prove the theorem (Singh, 1997, pp. 133–7). I cannot pretend to have any

great mathematical ability, but I do always try to remember to take with me a book of Sudoku puzzles when I fly, as an antidote to the boredom of the departure lounge and a distraction from the terrors of a bumpy flight. So there is nothing particularly special or unusual about responding to stress by choosing to manipulate one's mode of consciousness. However, when I focus on my Sudoku puzzle, I am aware of the fact that in doing so, I am choosing to shut out the uncomfortable situation in which I find myself. I might say, afterwards, that I had successfully generated the illusion that time had flown by more quickly. Once the flight is over, I can put my Sudoku puzzles aside. They are useful when I need to be distracted, but they are not weighty problems that demand a resolution. Boethius and Wolfskehl found consolation by turning to tasks that they considered important. In Boethius's case, there is the added bonus that the conclusion he arrives at – that God experiences all things simultaneously – implies that insofar as he has attained something approaching a timeless perception, freeing himself from the stresses of the line-mode, what he has attained is not a delusion, a refuge from reality, but something closer to the most accurate way of perceiving reality, the way it is perceived by God. Aquinas, too, affirms that God grants some people a share of his eternity, including the experience of unchanging thoughts (*Summa Theologiae*, Iae, Q. 10, Art. 11).

This, then, is what I mean by experience-enhancing theology, and it is important to note that I mean by this something more than theology that evokes certain experiences. Certainly, an essential feature of the *Consolations* is that we follow Boethius on a journey through different states of consciousness, each of which we are able to share by a sympathetic reading of his rich descriptions. However, as I have already noted, if this were all that were claimed for theology, it would seem to be merely an inferior substitute for music. What theology adds is an evaluation of the cognitive value of the states of consciousness we are going through, that is, an assessment as to whether they are helping us see reality from a better perspective, or drawing us in to potentially dangerous illusions. The intellectual element is not only part of the means by which Boethius achieves a different state of consciousness; having the backing of reason is necessary to generate a positive evaluation of his new state of mind.

4. Conclusion

At the start of this essay, I said that if theology is a complete waste of time, that is bad news for professional and amateur theologians alike.

The explanations for the existence of theology provided by cognitive scientists of religion (with the exception, perhaps, of Pyysiäinen) are really explanations for the existence of a professional class of religious experts who provide a service to the religious community and receive a reward (at the very least, a high status within the community) as a return. They are able to reassure believers that difficult questions can be solved (Dennett), and by ensuring a proper balance between A-system and B-system thinking, they maintain the stability of religious thought over time, providing a systematic ideological basis for religious communities (Whitehouse). These tasks may well be essential to the long-term health of a religious community, but they can safely be delegated to a few recognized experts who, by their investment of time and intellectual effort, demonstrate a high level of commitment to the community (Bulbulia). Experience-enhancing theology, on the other hand, is by its nature an amateur activity insofar as the benefits attained from practicing such a discipline accrue only to the practitioner. The term 'amateur' is not, of course, meant as a term of disparagement, nor do I mean that Boethius succeeded only in consoling himself. On the contrary, I am sure that through the centuries, many people have found comfort in reading the *Consolations* in time of crisis. However, some documents bring comfort by their mere presence. I keep a copy of my insurance policy in my car, and I am reassured by the knowledge that if I ever need my car to be fixed, I will be able to call upon a qualified person to do it. It is not written in my native language, and I have never felt the necessity of reading the small print. Its presence alone is a comfort. The *Consolations* is not like this: left to gather dust on the shelf it will provide no comfort until the day when the owner is ready to pick it up and read it.

One reason that cognitive scientists may have overlooked this form of amateur theology is that they tend not to be interested in altered states of consciousness, arguing that profound mystical experiences are not sufficiently widespread amongst religious believers to account for the prevalence of faith (Barrett, 2004, pp. 121–2). Indeed, it is somewhat refreshing that cognitive scientists explicitly raise the point that, for much of the time, the dominant state of mind for many participants in religious rituals is one of boredom (Whitehouse, 2004, pp. 97–9). However, the point about the paucity of mystical experiences, even if it is true, cannot be raised as an objection to the view I am proposing here. In the first place, I am not attempting to offer an explanation for the popularity of religious belief, which cognitive scientists are confident

they can explain, only for the theological beliefs of a minority, with which cognitive scientists seem to have more trouble. Second, the main goal of Donaldson's theory is not to explain highly unusual altered states of consciousness, although her work might have some relevance for such a study, but fluctuations in our state of mind that are a normal part of human development. It is no part of my argument that Boethius achieved some ineffable state of consciousness that is beyond the limits of normal human experience. Rather I am proposing that a state of mind that is not in itself unusual was given the seal of rational approval because of the content of his theology.

So, perhaps cognitive scientists of religion have something to learn from theologians about the motives for engaging in theology. If these motives remain obscure, cognitive scientists are liable to wonder whether theology is, in fact, a waste of time. The charge that this is the case is one that theologians must attempt to refute, and in attempting to meet this challenge, we must consider carefully the goals of theological activity, and the ways we have of recognizing whether those goals are being met. In other words, we must make explicit the standards by which we judge contributions to theology. My own proposal is that there are two sets of standards, which can, potentially, conflict. Technical theology aims as far as possible to preserve a set of intuitively attractive beliefs in the face of intellectual challenges. Experience-enhancing theology aims to provide an intellectual framework that reinforces some of the emotional resources that all of us can draw on to deal with the challenges of life. In this essay I have attempted to show how the existence of these two types of theology explains the persistent attraction of divine timelessness, despite the fact that this conflicts with popular religious ideas. Of course, to establish that this is an accurate model of theology would require further work, focusing on other divine attributes. That must be left for another occasion.

My concern in this essay has been to make explicit the reasons why anyone would want to believe in a timeless God rather than with the question of whether there is such a being. The best justification of time spent on theology would be to demonstrate that theology is capable of discovering truths. However, my goal was to look at implications of cognitive science of religion for the practice of theology, and I do not think that cognitive science can tell us anything one way or the other about the truth of religious beliefs. But if cognitive science does not offer a new source of truths about God, it can at least help theologians attain a higher level of self-knowledge.[5]

Notes

1. See Barrett, 2000 for a useful survey.
2. I take responsibility for this translation, although I make no claim to originality.
3. Donaldson provides a useful summary of the various modes in 1992, pp. 267–70.
4. The quotation appears in Donaldson, 1992, p. 155. It is taken from Ward, 1987/1998, p. 3. Donaldson cites Ward's book by its original title, *Images of Eternity*. It was republished as *Concepts of God*.
5. This essay is based on a paper with the same title that I presented at the conference on The Concept of God and the Cognitive Science of Religion, which took place at the University of Birmingham, June 2009. I would like to thank Florida State University – Panama's Faculty Development Committee for providing me with the funding to attend this conference, Yujin Nagasawa for organizing the conferences and the conference sponsors, which were the John Templeton Foundation and the Cognition, Religion and Theology Project at the University of Oxford. I received a lot of useful feedback, which has enabled me to improve the paper.

References

Aristotle (1999) *Nicomachean Ethics*, trans. Terence Irwin (Indianapolis: Hackett Publishing Company).

Barrett, J. (2000) 'Exploring the Natural Foundations of Religion', *Trends In Cognitive Science* 4.1: 29–34.

Barrett, J. (2004) *Why Would Anyone Believe in God?* (Walnut Creek, CA: Altamira Press).

Barrett, J., and F. Keil (1996) 'Conceptualizing a Nonnatural Entity: Anthropomorphism and God Concepts', *Cognitive Psychology* 31: 219–47.

Boethius (1990) *Consolatio Philosophiae*, ed. J. O'Donnell (Bryn Mawr, PA: Bryn Mawr College). Available online at www9.georgetown.edu/faculty/jod/boethius/boethius.html.

Boyer, P. (2001) *Religion Explained* (New York: Basic Books).

Bulbulia, J. (2009) 'Religiosity as Mental Time Travel: Cognitive Adaptations for Religious Behavior', in J. Schloss and M. Murray (eds), *The Believing Primate,* (New York: Oxford University Press), pp. 44–75.

Davidson, D. (2001) 'Mental Events', in idem, *Essays on Actions and Events* (Oxford: Clarendon Press), pp. 207–27.

Dennett, D. (2006) *Breaking The Spell* (London: Penguin).

Donaldson, M. (1992) *Human Minds* (London: Penguin).

Helm, P. (2001) 'Divine Timeless Eternity', in Gary Ganssle (ed.), *God and Time* (Downers Grove, IL: Inter Varsity Press), pp. 28–60.

Kenny, A. (1979) *The God of the Philosophers* (Oxford: Oxford University Press).

Marenbon, J. (2003) *Boethius* (New York: Oxford University Press).

Pike, N. (2002) *God and Timelessness* (Eugene, OR: Wipf & Stock).

Pyysiäinen, I. (2004) 'Intuitive and Explicit in Religious Thought', *Journal of Cognition and Culture* 4.1: 123–50.

Singh, S. (1997) *Fermat's Last Theorem* (London: Fourth Estate).
Slone, D. J. (2004) *Theological Incorrectness – Why Religious People Believe What They Shouldn't* (New York: Oxford University Press).
Stump, E., and N. Kretzmann (1981) 'Eternity', *Journal of Philosophy* 78: 429–57.
Ward, K. (1987) *Images of Eternity* (London: Darton, Longman and Todd); republished as (1988) *Concepts of God* (Oxford: Oneworld).
Whitehouse, H. (2004) *Modes of Religiosity* (Walnut Creek, CA: Altamira Press).

This page intentionally left blank

Part II
God, Creation and Evolution

This page intentionally left blank

3
Darwin's Argument from Evil

Paul Draper

The position that evolutionary biology makes the problem of evil worse for the theist is hardly new. Indeed, as we shall see, Darwin himself seems to have held it, and its more recent defenders include philosophers, scientists, historians and even theologians (though some of the evangelical Christian theologians who hold this position infer from it not that there is no God, but instead that the theory of evolution is false!). Consider, for example, the following three frequently quoted remarks. According to Bertrand Russell (1997, pp. 79–80),

> Religion in our day, has accommodated itself to the doctrine of evolution . . . . We are told that . . . evolution is the unfolding of an idea which has been in the mind of God throughout. It appears that during those ages . . . when animals were torturing each other with ferocious horns and agonizing stings, Omnipotence was quietly waiting for the ultimate emergence of man, with his still more exquisite powers of torture and his far more widely diffused cruelty. Why the Creator should have preferred to reach His goal by a process, instead of going straight to it, these modern theologians do not tell us.

Jacques Monod (1976) claims that

> [Natural] selection is the blindest, and most cruel way of evolving new species, and more and more complex and refined organisms . . . . The struggle for life and elimination of the weakest is a horrible process, against which our whole modern ethics revolts. An ideal society is a non-selective society, one where the weak is protected; which is exactly the reverse of the so-called natural law. I am surprised that a

> Christian would defend the idea that this is the process which God more or less set up in order to have evolution.

Finally, David Hull (1991, p. 486) maintains that

> The evolutionary process is rife with happenstance, contingency, incredible waste, death, pain and horror. . . . the God implied by evolutionary theory and the data of natural history . . . is . . . not a loving God who cares about His productions. He is . . . careless, wasteful, indifferent, almost diabolical. He is certainly not the sort of God to whom anyone would be inclined to pray.

These are rhetorically powerful statements, expressed with great conviction by highly respected thinkers; but at best they only hint at any sort of serious argument against the existence of God. Indeed, *fully developed* Darwinian arguments from evil (let alone convincing ones) are not easy to find. The politically correct stance nowadays, held by most atheists and almost all theists who accept evolution, is to deny the relevance of evolution to the issue of God's existence.[1] And some of those who dissent (e.g. Ayala, 2007) actually believe that evolution helps to solve or at least mitigate the problem of evil! Following a line of reasoning suggested by Darwin, I will dissent in the opposite direction by proposing a serious Darwinian argument from evil, an argument that is strongly suggested by Darwin's discussion of the problem of evil in his autobiography and elsewhere. My strategy (or, arguably, Darwin's strategy) is to use Darwin's theory of natural selection as a sort of 'atheodicy' – an explanation of a broad range of facts about good and evil that works much better when Darwin's theory is combined with a serious atheistic hypothesis than when it is combined with (orthodox) theism. I will begin by getting clear on some crucial terminology. Then I will offer an interpretation of Darwin's views about religion in general and about the problem of evil in particular. Finally, after identifying what is arguably Darwin's own argument from evil, I will try to develop and defend the argument, using some of Darwin's ideas, some ideas taken from contemporary evolutionary biology, and a few of my own ideas.

1. Terminology

By 'evolution' I mean the conjunction of two theses. The first asserts that evolution did in fact occur – complex life did evolve from simpler life. Specifically, it is the view that all multicellular organisms and

all (relatively) complex unicellular organisms on earth (both present and past) are the (more or less) gradually modified descendants of a small number of relatively simple unicellular organisms. The second addresses the issue of how evolution occurred. It states that almost all evolutionary change in populations of complex organisms occurs because of trans-generational genetic change or, more broadly, trans-generational change in nucleic acids. This thesis was not, of course, established until after Darwin's death, but it has since become an essential part of our understanding of evolution. While this second thesis is a claim about the mechanisms by which evolution takes place, it is important to distinguish it from the more specific claim that natural selection operating (indirectly) on random genetic mutation is the principal mechanism driving evolutionary change (or at least the principal mechanism driving the evolutionary change that results in increased complexity). It is this claim that I call 'the theory of natural selection' or sometimes just 'Darwin's theory', though Darwin of course did not know about the role of genetic mutation in producing the inheritable phenotypic traits that nature (directly) selects. In those parts of the essay in which I discuss specifically Darwin's own views, the phrase 'operating on random genetic mutation' can be understood to be replaced by 'operating on random variation'. Finally, I use the term 'Darwinian evolution' to refer to the conjunction of evolution and the theory of natural selection.

Darwin's main contribution to the theory of natural selection is the idea of natural selection. According to Darwin's understanding of this idea, evolutionary change results from competition for survival, random variation and heredity. A population of organisms will reproduce at a rate that guarantees a greater number of offspring than parents. This results in overpopulation, which is one of the main causes of competition for survival. This competition leads to evolutionary change because of variation and heredity. Although the characteristics of organisms in the population will vary only slightly and only randomly, such variation will suffice to give some organisms an advantage in the struggle for survival, thereby making these organisms more likely to reproduce. Because of heredity, these advantageous characteristics will often be passed on to offspring, and those offspring will also be more likely to reproduce. Thus, over generations the frequency of those characteristics in the population will gradually increase until eventually they become normal for the population. At this point, it is correct to say that the population has evolved. When enough of these small evolutionary changes accumulate over very long periods of time, the result is a new species.

According to Darwin, however, not all evolutionary change occurs because a characteristic confers an advantage in the struggle to survive. Some change occurs because a characteristic, though not itself advantageous, is correlated with characteristics that do confer such an advantage. Darwin called this 'correlated variation'. (Today it is associated with pleiotropy.) And still other change occurs, not because a characteristic makes an organism more likely to survive, but because it makes an organism more likely to reproduce, given that it survives long enough to do so. Darwin called this 'sexual selection'. Although Darwin restricted the term 'natural selection' to the first of these three mechanisms (which is sometimes called 'survival selection'), I will understand it to include all three (and others as well). Our contemporary understanding of natural selection is typically broadened to include a variety of causes, though not of course every cause, of differential reproduction.

2. Darwin on Religion

Darwin spent a fair amount of time reflecting on religion and especially on the religious implications of his theory of natural selection. In the part of his autobiography describing his views on religion and how they developed over time, he first briefly explains how he came to reject Christianity (1958a, pp. 71–2) and then begins a lengthier discussion of his reflections on the question of whether or not a personal (though non-Christian) God exists. For the most part, his focus is on the weaknesses of various arguments for God's existence, starting with William Paley's famous design argument. He had read Paley's *Natural Theology* (and two other books by Paley) while studying for a BA degree at Cambridge. Although Paley's argument previously seemed conclusive to him, he rejects the argument after discovering 'the law of natural selection', since that law shows that intelligent design is not necessary to account for the 'endless beautiful adaptations' in living things (1958a, p. 73). Interestingly, he distinguishes the issue of how to account for adaptations from the issue of whether 'the generally beneficent arrangement of the world [can] be accounted for' (ibid.). He claims, however, that natural selection can also account for this arrangement. I will return to his argument for this claim in section 5. After his discussion of the problem of evil, to which I will return momentarily, he rejects a number of other reasons for believing in God. He claims, however, that as late as the time he wrote *Origin of Species* he still could be considered a 'Theist' – 'Deist' would have been a better choice of words given his views on the problem of evil – because at that time he still accepted some sort of

cosmological argument from the existence of an 'immense and wonderful universe' to a 'First Cause having an intelligent mind in some degree analogous to that of man' (1958a, p. 77). He later came to believe, however, that 'The mystery of the beginning of all things is insoluble by us', on the basis of which he declares himself to be an 'Agnostic' (1958a, p. 78).

While Darwin's focus in the *Autobiography* is, as I said, on the absence of good reasons to believe in God, he also thought that there are good reasons to reject the omnipotent, omniscient and omnibenevolent God of orthodox theism. He was, without question, an atheist about that sort of God.[2] There is evidence of this both in his autobiography, especially as we shall see in the passage on the problem of evil, and also in his correspondence, especially his letters to the botanist (and orthodox theist) Asa Gray, who was a leading defender of Darwin's theory in the USA, but who wanted to reconcile it with teleology, divine design and even natural theology.

One of the most significant passages in the *Autobiography* (1958a, p. 75), especially for my purposes, is this one:

> That there is much suffering in the world no one disputes. Some have attempted to explain this in reference to man by imagining that it serves for his moral improvement. But the number of men in the world is as nothing when compared with that of all other sentient beings, and these often suffer greatly with no moral improvement. A being so powerful and so full of knowledge as a God who could create the universe, is to our finite minds omnipotent and omniscient, and it revolts our understanding to suppose that his benevolence is not [sic] unbounded, for what advantage can there be in the suffering of millions of the lower animals throughout almost endless time. This very old argument from the existence of suffering against the existence of an intelligent first cause seems to me a strong one; whereas, as just remarked, the presence of much suffering agrees well with the view that all organic beings have been developed through variation and natural selection.

This passage raises two questions. First, why does Darwin believe that theism's failure to explain suffering is an argument against theism? No theory (whether metaphysical or scientific) can explain even a small portion of the facts that need explanation. Thus, so long as the hypothesis that God exists is logically compatible with suffering, its failure to explain it appears no more significant than its failure to explain

the tides. Second, and even more importantly for my purposes, why does Darwin compare theism to the theory of natural selection here? Granted, that theory can explain many biological and psychological facts, including facts about pain and pleasure, better than theism can; but how is that supposed to strengthen the argument from evil against theism? Atomic theory can explain many chemical facts much better than theism, but that's no reason to doubt theism. After all, atomic theory can also explain many facts better than naturalism or deism or pantheism or any other *alternative* to theism that is not ad hoc.

3. A Preliminary Interpretation of Darwin's Argument from Evil

One possible answer to both of these questions is that Darwin holds that theism and the theory of natural selection are 'competing' or 'alternative' hypotheses in the sense that they are (known to be) mutually exclusive, at least on our background knowledge. (Two hypotheses are 'mutually exclusive on our background knowledge' just in case our background knowledge entails that at most one of them is true.) If Darwin believes this, then Darwin's answer both to the question of why theists need a theistic explanation of suffering and to the question of how his scientific theory of natural selection is relevant to the philosophical problem of evil is that his theory provides a good *atheistic* explanation of various facts about suffering (and pleasure). Chemists clearly cannot make any similar claim about atomic theory, which explains various chemical facts in a way that is, as far as we can tell, perfectly neutral between theism and atheism. A good atheistic explanation of what we know about suffering would threaten theism because, even though theistic belief is more often based on testimony (the testimony of one's parents in most cases) than on theism's explanatory power, theism's credibility is still challenged if a *serious* competing view, a competing view that is neither ad hoc nor initially less plausible than theism, explains certain facts better than theism does. Such a challenge would be a good prima facie reason to believe that the alternative view is more likely than theism to be true and so would be a good prima facie reason to believe that theism is probably false. While this is an interesting and plausible interpretation of Darwin's views, I intend to offer and defend a different interpretation for two reasons.

First, while some of Darwin's contemporaries regarded his theory as atheistic,[3] and while some of Darwin's remarks in various letters suggest that he believes that his theory is incompatible with the sort of

providence implied by orthodox theism, I doubt this interpretation can be correct simply because Darwin regards his view that natural selection drives most evolution as a highly confirmed theory, not as a mere hypothesis or conjecture. Thus, if he really believed that his theory entails atheism, then he would have no reason to appeal to an argument from suffering in order to defend his disbelief in beneficent design. He would simply infer the falsity of beneficent design and so of orthodox theism directly from the truth of his theory. Granted, he does reject Asa Gray's view that the (beneficial) variations selected by nature are 'intentionally and specially guided', but it is arguable that what he objects to here is the idea that God supplies these variations by miraculously intervening in nature, which would mean that the variations do not result from the operation of natural law. Since orthodox theism implies the existence of an omnipotent and omniscient God, it seems to allow for there being divine planning and remote causation (and so design) without direct divine action in nature. Darwin himself emphasizes that his theory is consistent with the view that all natural events result from designed laws. To be sure, he also favours the view that whatever good or ill results from the operation of these laws is a matter of chance, but that, I believe, is an inference he draws at least in part from the suffering in the world, not solely from the tenets of his theory.

This leads to or has already arrived at my second reason for rejecting any interpretation of Darwin's argument from evil that depends on Darwin taking his theory of natural selection (or its conjunction with background knowledge) to entail the falsity of orthodox theism. To interpret Darwin's theory this way is to attribute a false or at least poorly supported position to Darwin. The claim that most evolutionary change and most adaptations result from natural selection simply does not, so far as we can tell, entail atheism – not by itself, not when conjoined with what Darwin knew, not even when conjoined with current biological knowledge. This is controversial. Some thinkers, both at Darwin's time and today, believe that orthodox theism is incompatible with the randomness entailed by natural selection. Thus, before I move on to my preferred interpretation of Darwin's argument from evil, I will examine three arguments for such an incompatibility.

The first argument proceeds as follows. If the universe was created by a (temporal) omniscient being, then everything (with the possible exception of causally undetermined events) is foreknown. And if this omniscient being is also omnipotent and morally perfect, then anything affecting the well-being of sentient creatures was presumably not just foreknown but planned as well. But many particular aspects of nature,

including much of the variation in populations of organisms, affect the well-being of those organisms. Thus, orthodox theism entails that these aspects were planned, and so is incompatible with Darwin's theory, which entails that many of them were random and hence unplanned. This argument fails because it equivocates on the meaning of the word 'random'. Natural selection entails that useful variations in a population of organisms are random both in the sense that humans (because of our ignorance) cannot predict them and in the sense that any changes in the environment that make a particular variation useful do not directly cause that variation. But being random in these senses is consistent with being non-random in another sense. Specifically, as suggested above, it is consistent with being planned and indirectly caused by an omnipotent, omniscient and morally perfect person. If such a person created the universe and the natural laws that govern it, then that person might have planned, in accordance with those laws, exactly how the environment would change, exactly which variations would occur, exactly which traits would be selected for, and so on. Providence of this sort would not involve the sort of miraculous intervention in the developmental history of organic beings that would be incompatible with Darwin's theory.[4]

A second reason for thinking that orthodox theism is incompatible with the randomness entailed by the theory of natural selection is that some genetic mutations, for example mutations caused by x-ray irradiation, may depend on sub-atomic events governed by the probabilistic laws of quantum physics.[5] If, as most physicists believe, such events are causally undetermined, then some variations may be random in a stronger sense than even Darwin ever imagined. Thus, if we assume a temporal God and the logical impossibility of foreknowing causally undetermined events, then some of the variations upon which natural selection operates could not be planned even by an omnipotent and omniscient God. However, as long as a temporal person's being omniscient at some time only requires having as much knowledge as it is metaphysically possible to have at that time, this too is compatible (at least so far as we can tell) with orthodox theism. For an omniscient God would still know at any given time which future variations are physically possible, and so long as such a God had a providential plan for every contingency, no conflict with God's omnibenevolence is implied.

Finally, it might be argued that, because the variation involved in natural selection is random, natural selection is an inherently wasteful process, and it is this wastefulness that is incompatible with theism.

Now it is true that random variations are often 'useless' in the sense that they are not advantageous and may even be disadvantageous in the struggle to survive. It is also true that natural selection entails competition for survival and not all organisms can win such a competition. As others have pointed out, however, it is a mistake to claim that an omnipotent and omniscient person who produced such a process would be guilty specifically of 'waste'; for waste is possible only when one's resources are limited, and an omnipotent and omniscient being has unlimited resources. Of course, to the extent that such waste harms sentient beings, it may be incompatible with theism; but natural selection does not entail that such harm occurs nor does it entail that God has no good reason for causing or permitting that harm.

No doubt other possible reasons might be offered for believing that orthodox theism and the theory of natural selection are mutually exclusive given our background knowledge; but I am fairly confident that these will be no more successful than the three I have considered. If the literature on the logical problem of evil teaches us anything, it is that deriving specific claims about the world from theism is very difficult and so showing that theism is logically incompatible with our knowledge or theories about the world is also very difficult. Thus, if we want to avoid attributing to Darwin a false or at best poorly supported view about the relationship between his theory and orthodox theism, then we must reject this first interpretation of his argument from evil.

4. A Second Interpretation of Darwin's Argument

Fortunately, this doesn't return us to square one when it comes to interpreting Darwin's argument. We can keep the basic structure of the argument under the first interpretation, as long as we modify the argument's details enough to avoid the two objections to that interpretation but not so much that it loses all plausibility as an interpretation of what Darwin had in mind. Specifically, we can take as the hypothesis that competes with orthodox theism another view that Darwin defended, namely, that living things and their adaptations are not intentionally designed. Let's call this the 'no-design hypothesis'. Shortly before the passage in the *Autobiography* on the argument from suffering, he claims that 'There seems to be no more design in the variability of organic beings and in the action of natural selection, than in the course which the wind blows' (1958a, p. 73).[6] Further, he repeatedly (both in the *Autobiography* and elsewhere) connects his view that living things are not designed with examples involving suffering or death. For example, his

most famous comment about the problem of evil is in reality a comment about the absence of design in the world: 'I cannot persuade myself that a beneficent and omnipotent God would have *designedly* created the Ichneumonidae with the *express intention* of their feeding within the living bodies of caterpillars' (1958b, p. 249; my italics). What all this demonstrates is the plausibility of taking the no-design hypothesis to be the alternative to orthodox theism that Darwin implicitly has in mind when he discusses the problem of evil in the *Autobiography*.

The obvious advantage of this interpretation is that the no-design hypothesis, unlike the theory of natural selection, really is incompatible with orthodox theism. Of course, the purposes for which God designed living things may be unknown to us and Paley may be wrong that adaptations were designed in all cases for the purpose of benefiting the organisms that possess those adaptations; but if an omnipotent, omniscient, and omnibenevolent God created the universe, then the living world must be a providential one, and thus the ordered systems making up the living world are foreseen (at least as possible outcomes of an indeterministic process), ordained (at least relative to God's consequent will) and thus designed. Of course, the crucial question about the no-design hypothesis is whether it can explain any facts about good and evil better than orthodox theism. This is where Darwin's *theory* of natural selection comes into play on this interpretation. That theory can serve as a good 'atheodicy': an explanation of various facts about good and evil that works much better on the assumption that an alternative to theism – in this case the no-design hypothesis – is true than on the assumption that orthodox theism is true.

The disadvantage of this interpretation is that in the crucial passage about the problem of evil in his autobiography quoted earlier, Darwin doesn't mention the no-design hypothesis. He mentions his theory of natural selection. Recall the two questions about that passage discussed earlier. The first question was why Darwin thinks that theism needs to explain known facts about suffering. Both interpretations answer that question as follows: because a plausible atheistic hypothesis can explain those facts. The advantage of the second interpretation over the first is that it makes this answer true! The second question was why Darwin believes that his theory of natural selection is relevant to the problem of evil. The first interpretation can answer this question more easily than the second because it takes natural selection to be the crucial alternative hypothesis to orthodox theism. On the second interpretation, the answer is more complicated. Darwin's theory becomes, in effect, an auxiliary hypothesis that can successfully explain the suffering

(and pleasure) in the world when it is combined with the no-design hypothesis but not when it is combined with orthodox theism.

Of course, it is far from obvious that this answer to the second question is true. In the remainder of this chapter, I will offer some pro tanto reasons for thinking that it is true. I will divide my task into two parts. First, I will examine what facts about good and evil the theory of natural selection can explain. And second, I will argue that these explanations are much more successful when Darwin's theory is combined with the no-design hypothesis than when it is combined with orthodox theism.

5. Darwinian Explanations of Good and Evil

Let's begin with what Darwin says in the *Autobiography*. As I mentioned in section 2, Darwin thinks that his theory explains the 'generally beneficent arrangement of the world'. By the 'beneficent arrangement of the world', he means to refer to the alleged fact that, when it comes to the quantity of pleasure and suffering in the world, happiness 'decidedly prevails' over misery (1958a, pp. 73–4). He explains this alleged fact as follows:

> Every one who believes, as I do, that all the corporeal and mental organs (excepting those which are [not advantageous] to the possessor) of all beings have been developed through natural selection, or the survival of the fittest, together with use or habit, will admit that these organs have been formed so that their possessors may compete successfully with other beings, and thus increase their number. Now an animal may be led to pursue that course of action which is the most beneficial to the species by suffering, such as pain, hunger, thirst, and fear, – or by pleasure, as in eating and drinking and in the propagation of the species, &c. or by both means combined, as in the search for food. But pain or suffering of any kind, if long continued, causes depression and lessens the power of action; yet is well adapted to make a creature guard itself against any great or sudden evil. Pleasurable sensations, on the other hand, may be long continued without any depressing effect; on the contrary they stimulate the whole system to increased action. Hence it has come to pass that most or all sentient beings have been developed in such a manner through natural selection, that pleasurable sensations serve as their habitual guides. We see this in the pleasure from exertion, even occasionally from great exertion of the body or mind, – in the pleasure of our daily meals, and especially in the pleasure derived

> from sociability and from loving our families. The sum of such pleasures as these, which are habitual or frequently recurrent, give, as I can hardly doubt, to most sentient beings an excess of happiness over misery, although many occasionally suffer much. Such suffering, is quite compatible with the belief in Natural Selection, which is not perfect in its action, but tends only to render each species as successful as possible in the battle for life with other species, in wonderfully complex and changing circumstances. (1958a, pp. 74–5)

One problem here is that it's far from clear that pleasure really is predominant over suffering, and thus the alleged fact that Darwin is trying to explain may not actually obtain. Darwin himself was unsure. Though he says in the passage quoted above that he can 'hardly doubt' the predominance of pleasure, he admits elsewhere that it 'would be very difficult to prove' (1958a, p. 74). I believe that pleasure may be more common than pain in sentient beings that are flourishing (and notice too that Darwin's reasons for thinking the operation of natural selection would lead to the predominance of pleasure really only apply to beings that are flourishing). But a state of flourishing is not the norm for sentient beings in a Darwinian world, so Darwin's view is much too optimistic. We can, however, make use of some parts of his explanation of the alleged fact that pleasure predominates to explain the known fact that both pain and pleasure are used to promote reproductive success. Why not just suffering? Because excessive amounts of suffering tend to weaken an organism in a variety of ways. Why not just pleasure? Because pain works better when great or sudden dangers require a quick or extreme response. Of course, another advantage of being motivated by both suffering and pleasure, especially when it comes to biologically important activities like eating, is that one sort of motivation can still function if the other fails. For example, if an animal no longer derives pleasure from the taste of food, hunger may still motivate that animal to eat.

The theory of natural selection can also explain why sentient beings still suffer or feel pleasure even when, because of physical limitations or circumstances, such suffering and pleasure cannot causally contribute to survival or reproduction. For example, why does a person hopelessly trapped in a fire still feel pain? And why does a man still feel sexual pleasure after a vasectomy? If Darwin's theory is true, then one would not expect the mechanisms that produce pain and pleasure to be so fine-tuned as to eliminate such biologically useless pain and pleasure. For such fine-tuning would confer no advantage in the struggle for survival.

There are also a variety of more specific facts about suffering and pleasure that can be explained by the theory of natural selection, for example facts about the intensity of various sorts of suffering and pleasure. It can also explain a variety of facts indirectly, for example by explaining the anatomical causes of certain kinds of suffering. But instead of discussing these, which won't in the end make much difference to the strength of my argument, I would like to turn my attention to certain facts about goods and evils other than pleasure and suffering, specifically, the flourishing and languishing of sentient beings and the mixture of moral goodness and moral badness that we find in the human species.

Facts about flourishing and languishing are not reducible to facts about pleasure and suffering, since non-sentient organisms can flourish or languish and since a sentient organism might suffer much but still flourish, while another might, though this is much less likely, suffer little but still languish. Butler (1857, p. 253) wrote that

> of the numerous seeds of vegetables and bodies of animals, which are adapted and put in the way to improve to such a point or state of natural maturity and perfection, we do not see perhaps that one in a million actually does. Far the greatest part of them decay before they are improved to it; and appear to be absolutely destroyed.

I don't know how Butler arrived at his figure of less than one in a million, but he's certainly right that the vast majority of living things, including sentient beings, never flourish, many more flourish for only a very small portion of their lives, and almost none who live a full life flourish for all of it. A Darwinian explanation of this sad fact is readily available. If populations of organisms increase geometrically and this leads to competition for the resources necessary to survive, then inevitably a large percentage of all living things will not survive long enough to thrive, many more will barely survive and thus languish for all or almost all of their lives, and even those organisms that do flourish for much of their lives will, if they live long enough, ultimately languish in old age. A Darwinian world is inevitably cruel, especially to the young, the old and the genetically less fortunate.

Turning now to moral goods and evils, here we must proceed very cautiously. Moral phenomena are extremely diverse and it is very hard to separate out biological influences from cultural ones. The safest route is to identify those moral characteristics that humans share with at least some other animals, and then see if natural selection can explain these.

The most obvious is that we are basically self-centred: our tendency to behave in ways that promote our own welfare is typically much stronger than any tendency to act for the good of others. A Darwinian explanation of this is so obvious that it need not even be stated. But, like many other animals, we will at times sacrifice our own interests for the sake of others. Such altruistic behaviour can be divided into two types: kin altruism and non-kin or 'social' altruism.

Darwin himself was unable to offer any convincing explanation of how altruism of any sort arose. Thanks to the discovery of genes, however, we now have very plausible explanations of kin altruism. Because I share as many as half my genes with my kin, characteristics like self-centredness that help me survive and reproduce are not the only sorts of characteristics that can increase the likelihood that my genes will be passed down to future generations; characteristics like caring about my family members, which promote their survival and reproduction rather than my own, will also work. Thus, the theory of natural selection, when enriched by genetics, can account for kin altruism and the various natural virtues that produce it.

The origin of social altruism is tougher to explain. Since it is found only in humans and other higher mammals, perhaps it is a sort of by-product of kin altruism and intelligence. More likely, most social altruism is a form of reciprocal altruism, for which a variety of plausible Darwinian explanations have been offered. Social altruism is much weaker in most humans than kin altruism as well as more frequently absent altogether, and this is not surprising on the theory of natural selection. Notice also that it is typically very limited in its scope. The less like Smith some other sentient being is, the less likely it is that Smith will be concerned about that being's welfare. If, as Darwin thought, kin altruism is the basis for other sorts of altruism, then once again this is not surprising. Furthermore, it is clearly to be expected on Darwin's theory that universal altruism, that is, the tendency to sacrifice one's own interests for the sake of the interests of any sentient being, no matter how different from us and no matter how unlikely they are to reciprocate, is very rare in humans and usually very weak in those humans that have it. Having a strong dose of this characteristic would clearly not be advantageous in the struggle to survive. So its rarity, while unfortunate, is easily explained by the theory of natural selection. On the whole, then, the mixture of a basic self-centredness, with limited altruistic tendencies, can be explained quite plausibly by Darwin's theory.

Darwinian explanations have also been given for a variety of other more specific moral facts, like the greater frequency of marital infidelity

among the male of the species or the fact that child abuse is more common in step-parents. Some of these explanations are a bit far-fetched, but others are plausible despite their 'just-so' character. Of course, there are also some moral facts, like facts about self-destructive behaviour, that do not easily lend themselves to Darwinian explanations, but the theory of natural selection doesn't entail that every characteristic possessed by some members of a species (or even most members for that matter) must make reproduction more likely.

6. The Darwinian No-Design Hypothesis versus Darwinian Theism

The final and most important question that must be answered is why these Darwinian explanations of good and evil work much better when the theory of natural selection is combined with the no-design hypothesis than when it is combined with theism. I will offer three reasons. The first is that Darwin's theory is more likely on the no-design hypothesis than on theism. The second is that Darwinian explanations of good and evil are less complete when Darwin's theory is combined with theism than when it is combined with the no-design hypothesis. And the third is that, when combined with theism but not when combined with the no-design hypothesis, Darwin's theory mystifies certain facts.

There is much more disagreement about how probable the theory of natural selection is, either given the no-design hypothesis or given theism, than there is about how probable evolution is. The evidence for evolution is truly overwhelming. This is why almost all well-educated theists are also evolutionists. But Darwin's theory is another matter entirely. Few deny that natural selection accounts for some evolutionary change, but, according to the *theory* of natural selection, natural selection operating on random genetic mutation is the *principal* mechanism driving evolution. And many well-educated theistic evolutionists are much less enthusiastic than their atheistic counterparts about this claim. Not that there is any consensus among theistic evolutionists. Some believe that Darwin's theory is highly probable on both the no-design hypothesis and on theism; others believe it is very improbable on both the no-design hypothesis and theism; and others think it more probable on theism than on the no-design hypothesis (because, they claim, it could only work with divine assistance).

Still others agree with my position, which is that the theory of natural selection is highly probable on the no-design hypothesis but not on theism. This position entails that naturalistically inclined evolutionists who

have great confidence in the truth of the theory of natural selection and theistic evolutionists who have serious doubts about it are both right, *conditional on their respective metaphysical positions*. This position is based in part on my belief that the scientific case for the theory of natural selection is impressive but far from complete. It is incomplete because we don't know in any detail how most of the characteristics of living things evolved by natural selection. Of course, if Darwin's theory is true, we wouldn't expect to know all the details. And no one has shown that any instance of adaptive complexity in present-day living things *cannot* be accounted for by natural selection broadly understood. So these sorts of gaps in our knowledge don't refute the theory of natural selection. Nor do they prove that natural selection can work only with divine assistance of some sort. But they do make the case for the theory incomplete.

If, however, one assumes that the no-design hypothesis is true, then this incomplete case suffices to make the theory of natural selection highly probable because there are no viable naturalistic alternatives to it. Of course, one hears about scientific challenges to Darwinian evolution. But the only plausible ones are either challenges to strict gradualism or attempts to show that natural selection is not the only mechanism driving evolution. The theory of natural selection, as I define it, does not entail strict gradualism; so punctuated equilibriumism is compatible with it. Nor does it entail that all evolutionary change is caused by natural selection; so genetic drift, symbiosis and other mechanisms may also be at work, especially when the change in question does not involve a substantial increase in complexity. Thus, for the proponent of the no-design hypothesis, the theory of natural selection as I *broadly* define it really is the only game in town.

On theism, however, there are viable alternatives, and this means that the gaps in our knowledge mentioned above make it unreasonable for a theist to be confident in the truth of the theory of natural selection. What are these alternatives? It's hard to know where to begin. For if the universe was created by an omnipotent and omniscient being, then so many logical possibilities are also *real* possibilities! For example, if theism is true, then it is a real possibility that God has on billions of occasions in the history of the development of life on earth miraculously caused various mutations for the purpose of modifying a species in ways that further God's goals. Another real possibility is that most traits that arose by random variation have been artificially selected by God rather than winning a fair competition in the struggle for life. Or perhaps God never directly intervenes, but there is an as yet undiscovered hidden mechanism that has caused most of the evolutionary change that has

taken place. This last scenario is, of course, also possible on the no-design hypothesis, but on theism it is not just possible but *plausible*. So one reason that Darwinian explanations of good and evil work better for defenders of the no-design hypothesis than for theists is that the theory of natural selection is much more probable on the no-design hypothesis than on theism.[7]

A second reason is that Darwinian explanations of facts about good or evil will be less complete if theism is true than if the no-design hypothesis is true. To see why, I will borrow a couple of definitions from Richard Swinburne. Swinburne (2004) defines a *full* explanation as follows: 'An explanation of E by F is a full one if F includes both a cause, C, and a reason, R, which together necessitated the occurrence of E' (p. 76). A *complete* explanation is a special kind of full explanation. F is a complete explanation of E if it is a full explanation and none of the factors it cites can be explained by any contemporaneous factors (p. 78). The key word here is 'contemporaneous'. A complete explanation need not be an ultimate explanation. It need not be such that nothing prior to it explains the factors it cites.

Darwinian explanations of good and evil are obviously not ultimate explanations. And even if the no-design hypothesis is true, they are not full since the factors they cite do not necessitate the facts they explain. My claim, however, is that they are 'less complete' if theism is true, by which I mean that they are not as close to being complete explanations. This is so, because, if theism is true, then God is omnipotent and hence the existence of the universe depends on God at any given time. God, in other words, did not just create the universe, he sustains it. But that means any complete explanation of facts about good and evil must, if theism is true, include God's moral justification for allowing those facts to obtain.

To put the point another way, Darwin's theory comes closer to solving the puzzle of good and evil faced by the proponent of the no-design hypothesis than the puzzle of good and evil faced by the theist. If the no-design hypothesis is true, then, the causes of good and evil are axiologically indifferent, and thus a naturalist will want to know why good and evil are so abundant. Darwinian explanations, though not ultimate explanations, enhance a naturalist worldview by providing intellectually satisfying explanations consonant with that worldview. Fill in some more details and those explanations could be complete. The theory of natural selection, however, does little to solve the theist's perplexity about good and evil. For God wouldn't allow natural selection to determine the pattern of good and evil observed in the biological world

unless he had good moral reasons for doing so. And given the failure of existing theodicies, we have no idea what those good reasons are. For the theist, Darwinian explanations leave the biggest questions about good and evil unanswered.

A third way of stating essentially the same point is that, whenever a theist uses a Darwinian explanation of facts about good and evil, he or she is in effect implicitly adopting a number of completely ad hoc and very specific auxiliary hypotheses. If the no-design hypothesis is true, no such auxiliary hypotheses are needed or presupposed. For example, is there some greater good that, because of its logical connections to suffering, requires that suffering be used to motivate animals to pursue the biological goal of self-preservation? Does some moral end makes it desirable for suffering to continue even when it serves no biological purpose? Is there a moral justification for using natural selection to produce a world in which most living things never or rarely flourish because they compete with each other for survival? Or to produce a species that possesses as little natural virtue and as much natural vice as humans do? On the no-design hypothesis, Darwinian explanations of good and evil work whether the answer to these and many similar questions is 'yes' or 'no'. But on theism, the answer must in every single case be 'yes'. On theism, natural selection cannot drive evolution unless its doing so coincides with the moral goals of an all-powerful, all-knowing and morally perfect creator. And that's a really big coincidence that the no-design hypothesis doesn't need.

A third reason that Darwinian explanations of good and evil work better on the no-design hypothesis than on theism is that the theory of natural selection, when combined with theism, mystifies certain facts. Mystification is the opposite of explanation. A hypothesis 'mystifies' a fact if it makes that fact more puzzling rather than less puzzling – if it raises why-questions about that fact rather than answering them.

On Darwinian theism, harm to living things is not an unfortunate side effect of a good creation, but rather is an integral part of the process, started long before any human beings existed, by which God chose to create various living things. Thus, if Darwin's theory is true, then harm to living things is not, for example, the result of human beings or other free agents rebelling against God and acting contrary to his purposes. Rather, God has purposely used such harm as a means of achieving his creative purposes. And this would require a very strong sort of moral justification. Thus, the theory of natural selection makes the evil in nature even more shocking for theists than it would otherwise be.

Another problem that's exacerbated by adding Darwin's theory to theism is the problem of animal suffering. Theists have for centuries tried

to avoid this problem by at least quasi-Cartesian means. While almost no theists today deny, as Descartes did, the reality of animal suffering, they still try to minimize its significance or justify the fact that their theodicies ignore it by claiming that we know very little about what the mental states of animals are like, that for all we know animal suffering isn't much like human suffering and may even be of no moral significance.[8] If, however, Darwinian evolution is true, then we are the more or less *gradually modified* descendants of (non-human) animals. This means that, when humans and closely related animals behave very similarly in very similar circumstances, it is likely that they share very similar mental states. Thus, for example, when animals exhibit what overly cautious (or overly Skinnerian) psychologists might describe as a 'fear-like response' in circumstances that would cause fear in humans, it is likely that what those animals feel is very similar to human fear. In general, then, it is likely on Darwinian evolution that higher mammals suffer in many of the same ways and just as intensely as we do.[9]

So animal suffering is as real and as serious a problem for Darwinian theists as human suffering. In fact, to make matters worse, Darwinian theists are forced to admit that animal suffering is similar to human suffering *despite the fact that animals are not moral agents*. This deals a severe if not deadly blow to the otherwise plausible position that, if God exists, then our being moral agents has something to do with why God allows us to suffer to the extent we do. Thus, it is not just animal suffering that is mystified by mixing the theory of natural selection with theism. Human suffering becomes more puzzling, too.

7. Implications

We have, then, three very strong reasons for believing that the marriage of Darwinian biology to theistic metaphysics is an unhappy one. Although theism does not entail that Darwinian explanations of good and evil are false, it does greatly weaken them. It weakens them because natural selection is unlikely on theism, because ad hoc and specific auxiliary hypotheses are presupposed, and because other facts are mystified. This would not be a problem for orthodox theism if powerful theodicies were available, but no such theodicies exist. The few theodicies that do offer decent explanations of some facts about evil fail to offset the force of Darwin's argument from evil for two reasons. First, they adequately account for relatively few facts about suffering or other evils. Second, their ability to explain the few facts they do explain is purchased at the price of making other facts even more mysterious. For example, if we assume that God wants to use suffering to build moral character, then

certain sorts of suffering and other conditions that are, quite predictably, demoralizing become even more puzzling. Or, if it is free will that is so important, then why are so many human beings robbed by a variety of (natural) circumstances of any genuine freedom?

If, instead of focusing on a few isolated cases, one looks at the overall pattern of evil in the world, being careful not to ignore the non-human part of that world, then one cannot help but be struck by its apparent moral randomness. The theory of natural selection, when combined with the no-design hypothesis, offers a systematic explanation of a broad range of facts about good and evil with which no theodicy I know of can even begin to compete. Therefore, since the theist has no decent and systematic explanation of these facts – or at least none that is likely to be true if theism is true – while the proponent of the no-design hypothesis does, it follows that these facts are powerful evidence supporting the no-design hypothesis over theism. Further, the no-design hypothesis is both initially plausible and fits with much background knowledge better than theism. An evolutionary process that, instead of being beneficently designed, is not designed at all is more likely to take a very long time, to include imperfect adaptations, and to involve adaptations that are harmful for individuals (even when they are good for the genes those individuals host). And this is, in fact, what we observe. Therefore, in the absence of some other evidence supporting orthodox theism over the no-design hypothesis, it follows that the no-design hypothesis is much more probable than theism and hence, other evidence held equal, theism is very probably false.[10]

Notes

1. I use the term 'politically correct' here because of the broadly political motivations that many have for holding that evolution is religiously neutral. These motivations include the desire to prevent creationism or intelligent design theory (ID) from being taught in science classes in public schools in the USA. In his ruling against the constitutionality of teaching ID, Judge John Jones (2005, p. 136) of the US District Court in Pennsylvania based his ruling in part on an appeal to the view of 'scientific experts' who 'testified that the theory of evolution . . . in no way conflicts with, nor does it deny, the existence of a divine creator'.
2. Darwin consistently denied being an atheist. For example, in a letter to one J. Fordyee, dated 1879, he said, 'In my most extreme fluctuations I have never been an Atheist in the sense of denying the existence of God' (1958b, p. 59). There is good reason to believe, however, that 'God' is being used here in a broad sense. It is telling, for example, that when a German student inquired in the very same year about Darwin's religious views, a family member wrote back that Darwin 'considers that the theory of Evolution is

quite compatible with the belief in a God; but that you must remember that different persons have different definitions of what they mean by God' (1958b, p. 61).

3. For example, Charles Hodge (1874, pp. 176–7) wrote in a book called *What is Darwinism?*, 'We have thus arrived at the answer to our question. What is Darwinism? It is atheism. This does not mean . . . that Mr Darwin himself and all who adopt his views are atheists; but it means that his theory is atheistic; that the exclusion of design from nature . . . is tantamount to atheism.'
4. Darwin (1958b, p. 249) himself appears to affirm this possible reconciliation of his theory with divine design in one of his letters to Asa Gray.
5. Philip Kitcher (1982, p. 87) makes this point.
6. One of Darwin's favourite arguments for the conclusion that God does not guide natural selection by directly causing certain variations is found (among other places) at the end of his book on *The Variation of Animals and Plants under Domestication*, as Francis Darwin notes in his *Autobiography* (1958a, p. 73 n. 60). He says there that, since it is impossible to believe that God intentionally provides variations for the sake of breeders, it is unreasonable to believe that he intentionally provides the variations used in natural selection. Although Darwin claims that this argument has never been answered (1958a, p. 73), it isn't difficult to provide such an answer. Given the less than respectable goals of many breeders, I agree that it would be unreasonable to believe that all of the variations used by breeders are intentionally provided by a morally perfect God for the benefit of the breeders. But from the fact that they were not intentionally provided *for the benefit of the breeders*, it does not follow that they were not intentionally and specially provided, and hence it provides no reason at all for thinking that the variations used in natural selection are not in at least some cases produced by God's miraculous intervention.
7. It is worth noting that, if the theory of natural selection someday becomes virtually certain on both the no-design hypothesis and theism, then the truth of that theory would itself be evidence favouring the no-design hypothesis over theism because natural selection would, for the reasons mentioned above, still be *antecedently* more probable on the no-design hypothesis than on theism.
8. See, for example, Murray, 2008.
9. For an excellent statement and defence of this sort of argument, see Rachels, 1991, ch. 4.
10. Sincere thanks to William Hasker, Yujin Nagasawa and Beth Seacord for helpful comments on earlier drafts. I am also grateful for generous support from the Center for Philosophy of Religion at the University of Notre Dame, the John Templeton Foundation and Purdue University.

References

Ayala, F. (2007) *Darwin's Gift: To Science and Religion* (Washington, DC: Joseph Henry Press).

Butler, J. (1857) *The Analogy of Religion, to the Constitution and Course of Nature* (Philadelphia: J B Lippincott & Co.).

Darwin, C. (1958a) *The Autobiography of Charles Darwin, 1809–1882*, ed. N. Barlow (New York: W.W. Norton & Co.).
Darwin, C. (1958b) *The Autobiography of Charles Darwin and Selected Letters*, ed. F. Darwin (New York: Dover Publications).
Hodge, C. (1874) *What Is Darwinism?* (New York: Scribner, Armstrong).
Hull, D. (1991) 'The God of the Galapagos', *Nature* 352 (8 August): 485–6.
Jones, J. (2005) 'Memorandum Opinion', Kitzmiller et al. v. Dover Area School District et al. Available at www.pamd.uscourts.gov/kitzmiller/kitzmiller_342.pdf.
Kitcher, P. (1982) *Abusing Science: The Case against Creationism* (Cambridge, MA: The MIT Press).
Monod, J. (1976) 'The Secret of Life', interview with Laurie John, Australian Broadcasting Co. (10 June).
Murray, M. (2008) *Nature Red in Tooth and Claw: Theism and the Problem of Animal Suffering* (New York: Oxford University Press).
Rachels, J. (1991) *Created from Animals: The Moral Implications of Darwinism* (New York: Oxford University Press).
Russell, B. (1997) *Religion and Science* (New York: Oxford University Press).
Swinburne, R. (2004) *The Existence of God*, 2nd edn (Oxford: Oxford University Press).

4

Attributing Agency: Fast and Frugal or All Things Considered?

Graham Wood

1. Introduction

We attribute agency to ourselves and other people. We also attribute agency to other physical entities or systems, for example, our cat, computer or car. We might also attribute agency more broadly: to a river, an ecosystem or the universe as a whole. The attribution of agency in this broader sense is central to religious belief, and is often characterized as the assertion of the existence of gods and spirits. For example, Australian Indigenous peoples believe that spirits formed the landscape, Lowland Maya people hold beliefs about forest spirits, and Christians, Muslims and Jews believe in the existence of the god of classical monotheism. All these peoples are attributing agency as a central feature of their religion. This chapter explores this attribution of agency. Using insights from contemporary cognitive science and philosophy of mind, I offer a characterization of the cognitive process by which agency is attributed to ourselves, and religious entities. Along the way I address two questions. How do we attribute agency? And, why do we attribute agency?

The characterization I offer adapts Dennett's intentional stance (1987). I suggest that the intentional stance is a good way to understand the attribution of agency to humans. However, I further suggest that with one adaptation it can also be used to understand the attribution of religious agency. I begin the chapter with a discussion of several interrelated themes that inform the position I am advancing: evolutionary psychology, the dual process theory of mind and the modularity theory of mind. And, in the light of these themes, I consider the three stances of Dennett: the physical, design and intentional stances. I then present an adaptation of Dennett's intentional stance and use it to characterize

the attribution of agency in religious contexts, what I call 'taking the spiritual stance'.

Now, for Dennett, if one is able accurately to predict the behaviour of a particular physical system (for example, a human body), by attributing beliefs and desires to it, then that physical system *has* beliefs and desires (i.e. intentional states) in what he calls a 'semi-realist' sense. In other words, the test for whether a system has beliefs and desires is whether one is able successfully to predict the behaviour of that system by attributing beliefs and desires to it. Thus, the test is predictive success. If a system passes this test, then the system has beliefs and desires, in Dennett's semi-real sense. He calls these beliefs and desires 'real patterns'. Using my adaptation of the intentional stance, I argue that the beliefs and desires of religious entities can also be understood as 'real patterns' in Dennett's semi-realist sense. The position I advance relies on a conceptual relationship between the attribution of agency and the attribution of beliefs and desires to a physical system as follows. If beliefs and desires are attributed to a system, then agency is attributed to that system.

Now for the record, Dennett's intentional stance and the conceptualization of beliefs and desires as real patterns is not the theory of mind standardly endorsed by contemporary cognitive science. Dennett's theory of mind can be understood as instrumentalist. Instrumentalist theories endorse concepts (such as beliefs and desires) because they are *useful* in predicting events (such as the behaviour of humans). Notice how beliefs and desires get endorsed as 'real patterns' on Dennett's account *because* they are predictively successful.

My characterization of the attribution of beliefs and desires in the religious context is also instrumentalist. But it relies on a broader understanding of instrumentalism, as it is understood within the tradition of classical philosophical pragmatism (James, 1975). And in particular I rely on this broader notion of instrumentalism to justify my adaptation of Dennett's intentional stance. My account of the attribution of agency within religion is a pragmatist account. Philosophical pragmatists, among other things, take concepts to function as instruments (Hookway, 2010). To understand this approach to concepts, consider William James on the concept of truth: 'The true is the name of whatever proves itself to be good in the way of belief, and good, too, for definite assignable reasons' (1975, p. 42). '*"The true", to put it very briefly, is only the expedient in the way of our thinking* [. . .] Expedient in almost any fashion; and expedient in the long run and on the whole, of course' (1975, p. 106; emphasis original). Notice that on both characterizations

of the true above, James includes conditions. On one account, the true must be good for 'definite assignable reasons.' And on the other, the true must be 'expedient in the long run and on the whole.' Here there is a parallel with the instrumentalism of Dennett's intentional stance. Dennett's instrumentalism can be understood as a version of James's pragmatism by simply stipulating that *predictive success* is the 'assignable reason' or the relevant 'expediency'. However, this prompts the question: Why should *predictive success* be the relevant expediency? In my adaptation I change the 'expediency', or the 'definite assignable reason', to *reproductive success over evolutionary time*.

In Dennett's original intentional stance, the test is predictive success. If a system passes this test, then the system has beliefs and desires, in Dennett's semi-real sense. My adaptation changes the test from 'predictive success' to 'reproductive success over evolutionary time'. Using the new test, if the attribution of beliefs and desires to a system leads to the reproductive success (of the lineage) of the organism making the attribution (over evolutionary time), then the system has beliefs and desires. This adaptation is justified as follows. From an evolutionary perspective the 'function' of cognitive systems is not, first and foremost, accurate prediction, but rather it is to keep the organism reproducing. Accurate prediction has a role in reproductive success, but it is not the foundational role. I suggest that the attribution of agency within religious belief has led to reproductive success over evolutionary time for the lineage *Homo sapiens*. And thus my account helps explain *why* we believe in gods and spirits. But, as I noted above, this chapter also addresses the question of *how* we attribute beliefs and desires. Thus, I will also explore the *cognitive origins* of Dennett's 'semi realist' beliefs and desires.

Dennett describes beliefs and desires as real patterns. But for Dennett the perception of these patterns is fundamentally dependent upon the organism doing the attributing. I take this to imply that the attributing organism itself to some extent *constructs* the beliefs and desires attributed to other systems. And taking this a little further, I suggest that the concept of a 'belief' or 'desire' is an *heuristic* concept produced by a module within the mind of the organism doing the attributing. And once beliefs and desires are attributed to a system, a natural way to think of this system is *as an agent*, thus agency is also attributed to the system. Borrowing the distinction from the heuristics literature between 'fast and frugal' thought and 'all things considered' thought (Gigerenzer et al. 1999), I suggest, the concept of agency is not an 'all things considered' concept, but rather a 'fast and frugal' *heuristic*

concept. Thinking in terms of agency is a way of thinking that increases the reproductive success of organisms that employ the method, and so is a way of thinking that has been retained over the evolutionary time.

The chapter ends with some speculation concerning how the attribution of agency in religious contexts has furthered the reproductive success of our species over evolutionary time. And so, now I begin the substantive content of the chapter by discussing the assumptions I make and the perspective I take throughout.

2. Background Themes and Assumptions

Evolutionary psychologists assume that the mind is an evolved system, and use this to guide their choice of theories about the structure and functioning of the mind. 'Evolutionary analysis' (Samuels, 2000, p. 24) imposes constraints on the type of theory it is reasonable to accept. If a theory proposes a structure and function of the mind that evolution is very unlikely to have produced, then that counts against the theory. Alternatively, if a theory proposes a structure and function of the mind that evolution is very likely to have produced, then that counts in favour of the theory. So this leads to the question: What structures and functions is evolution likely to have produced? This question can be answered by looking at other evolved structures. The structure and function of other organs in the body all have an important similarity. *They are specific solutions to particular survival problems*. For example, the heart pumps blood, it does not oxygenate or filter the blood, and the kidneys filter the blood, they do not pump or oxygenate it. The lungs oxygenate the blood, they do not pump or filter it. So evolution produces specific solutions to specific survival challenges, rather than general solutions to general survival challenges. This *evolutionary analysis* leads inductively to the conclusion that cognitive systems are also specific solutions to specific problems faced by our hominid ancestors (Samuels et al., 1999, pp. 85–6).

The lesson evolutionary psychologists take from this analysis is that it is wrong to think of the mind as *one* general information processing system. Rather, the mind should be thought of as a *set* of distinct information processing systems (Carruthers, 2006, pp. 12–28, cf. Simon, 1962). Now, evolutionary psychology conflicts with orthodox psychology because the latter assumes that a set of *general cognitive capacities*, including habituation, operant and classical conditioning, imitation and the rules of logic and probabilistic reasoning, are responsible for most of our cognitive abilities. However, the perspective of evolutionary

psychology is gaining ground among both psychologists and biologists, as illustrated by the observation that psychologists 'should be trying to identify and understand these specialized circuits rather than pretending that human behaviour can be derived from a few law-like mechanistic principles' (Wilson, 2002, p. 26). If we assume that the mind is the result of evolution, then, it is argued, we should think about the mind differently. For example, consider the dual process hypothesis and the modularity hypothesis.

I assume that the mind has (at least) two distinct modes of thought. This has been called the *dual process theory* or *two systems hypothesis*. I will not argue for this, but rather simply assume it in the discussion that follows. Stanovich and West (2000) called these two systems System 1 and System 2. They are characterized as follows:

> *System 1 (intuition)*: fast; automatic; undemanding of cognitive capacity; acquired by biology, exposure, and personal experience.
>
> *System 2 (reasoning)*: slow; controlled; demanding of cognitive capacity; acquired by cultural and formal tuition. (Kahneman, 2002)

Kahneman (2002) characterizes System 1 as generating intuitive judgements that 'occupy a position – perhaps corresponding to evolutionary history – between the automatic operations of perception and the deliberate operations of reasoning' – leaving the deliberate reasoning to System 2.

Another way to conceptualize the distinction drawn between System 1 and 2 is to use the notion of modularity. Fodor (1983) first proposed the notion of mental modules. Fodorian modules (part of System 1) are said to receive information coming into the mind, via, for example, the sense organs, and manipulate this information before it is then made available to consciousness (System 2). A standard example of a Fodorian module at work is the process that leads to the Müller–Lyer illusion.

In this illusion, involving two lines of equal length with chevrons pointing in different directions on each of the two lines, the conscious

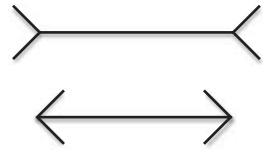


Figure 4.1 The Müller–Lyer Illusion

mind cannot help but see the lines as unequal (cf. Segall et al., 1966). Given that the lines are in fact the same length, presumably the information received by the retina corresponds to lines of equal length. It is assumed that this illusion is the result of a manipulation of the information stream between the retina and the conscious mind such that the lines of equal length are represented to consciousness as lines of unequal length. And it is assumed that a module is responsible for this manipulation.

For the purposes of this chapter it is important to note two features of Fodorian modules: first, the outputs they produce are mandatory, and second, the manipulations made inside modules are relatively inaccessible to System 2. Fodor describes these as follows:

> It is worth distinguishing the claim that input operations are mandatory (you can't but hear an utterance of a sentence as an utterance of a sentence) from the claim that what might be called 'interlevels' of input representation are, typically, relatively inaccessible to consciousness. Not only must you hear an utterance of a sentence as such, but, to a first approximation, you can hear it only that way. (Fodor, 1983, p. 55)

In the case of the Müller–Lyer illusion, we must see the lines as having different lengths, and we cannot see them any other way. If we were able to see them as both the same length and different lengths at the same time then our visual image would be incoherent.

While Fodor believed that modules existed only on the periphery of the mind, others have suggested that the mind is massively modular. For example, Cosmides and Tooby hold that many (non-peripheral) cognitive processes are modular:

> our cognitive architecture resembles a confederation of hundreds or thousands of functionally dedicated computers (often called modules) designed to solve adaptive problems endemic to our hunter-gatherer ancestors. Each of these devices has its own agenda and imposes its own exotic organization on different fragments of the world. There are specialized systems for grammar induction, for face recognition, for dead reckoning, for construing objects and for recognizing emotions from the face. There are mechanisms to detect animacy, eye direction, and cheating. There is a 'theory of mind' module . . . a variety of social inference modules . . . and a multitude of other elegant machines. (Cosmides and Tooby, 1995, pp. xiii–xiv)

The modules that are said to exist in the massive modularity hypothesis are not all Fodorian modules. For example, some modules may be less mandatory, and central cognition may have more access to the internal workings of those modules. But the concept of a modular system *that is to a significant extent independent from the operation of central cognition* will be central to this chapter. And I assume that while System 2 is concerned with the 'deliberate operations of reason', System 1 processes are largely independent processes undertaken by modules in the mind. Furthermore, I suggest that, to a large extent, *System 1 actually provides System 2 with the concepts with which it thinks*. In other words, while System 2 might be responsible for putting all the pieces of the jigsaw (that is, conscious thought) together, it is System 1 that provides System 2 with the pieces of the jigsaw. To illustrate this with a famous example, System 1 may provide System 2 with certain pieces of the jigsaw, for example, 'I' and 'am', and then System 2 puts them together: 'I think therefore I am.'

Two relevant modules are the 'folk physics' module (Spelke, 1990) and the 'folk psychology', or 'theory of mind' module (Baron-Cohen, 1995). However, while in the original presentation each of these was taken to be a single module, I think of these as a set of modules (cf. Gerrans, 2002). So I assume a set of modules generates the contents of both 'folk physics' and 'folk psychology'. On the modular account of 'folk physics' there is a dedicated part of the mind (a module, or set of modules) that processes information about physical objects or systems. This module or set of modules allows us quickly and accurately to predict the behaviour of physical objects, so we can successfully interact with those objects. On the modular account of 'folk psychology' there is a dedicated part of the mind (again, a module, or set of modules) that processes information about objects or systems that are taken to have intentional states, roughly speaking, objects or systems that are taken to have beliefs or desires, or agency (for example, humans). This module or set of modules allows us efficiently to interact with these intentional systems.

3. System 1 Modules Generate Heuristic Concepts

Earlier I suggested that 'beliefs' and 'desires' can be understood as 'heuristic concepts'. I will now explore this idea further. System 1 produces thought that is fast, automatic and undemanding of cognitive capacity. This characterization has much in common with the notion of heuristic thought (Gigerenzer et al., 1999; Gilovich et al., 2002). An heuristic is a 'rule or solution adopted to reduce the complexity of

computational tasks, thereby reducing demands on resources, such as time, memory, and attention' (Richardson, 1999, p. 379).

An example is the recognition heuristic. If you are not a German and you are asked the following question, 'Which of Dortmund or Munich is the larger city?', you may use the fact that you recognize the name of one city and not the other as a fast and frugal way of deciding that Munich was the larger city (Gigerenzer et al., 1999, p. 41). This works because there is a correlation between the population of a city (the target attribute) and the likelihood that you recognize the name of the city (the heuristic attribute). The fact that one can consciously reflect on the way this heuristic works shows that it is not modular, in the sense that the process is inaccessible to conscious awareness, thus it is employed by System 2 (deliberative reasoning).

But, I suggest, System 1 employs similar heuristic techniques (within modules). And if the modular process is mandatory and the internal workings are inaccessible to System 2, then the conscious awareness of such an heuristic process would be very different to our experience of using the recognition heuristic, say. It might seem more like our experience of the Müller–Lyer illusion. In the case of the Müller–Lyer illusion it is assumed the module uses the opposite directions of the chevrons as visual cues to indicate the relative distance of the two lines from the viewer. The module (erroneously) interprets one line as closer than the other and thus manipulates the lengths of the lines before presenting the image of the lines to consciousness. So, the fact that the lines appear as having different lengths is itself an heuristic 'perception' that System 1 has provided to System 2 to allow System 2 quickly to make choices about the relative distance of objects.

And if we take massive modularity seriously then we should expect to find other modules manipulating information. Furthermore if these other modules are mandatory and the processes occurring within the module are inaccessible to System 2, then again, just as is the case with the Müller–Lyer illusion, we need to be more careful about accepting the veridical nature of the outputs of these modules. Take for example the 'folk physics' module (or set of modules), which is said to provide consciousness with information about how mid-sized physical objects move.

As we learn more about the process by which visual information is manipulated by the mind we find that there are specific parts (modules?) of the visual cognition system that are responsible for edge perception, and other parts (yet other modules?) that take perceptions of edges as

inputs and then manipulate this information to generate the perception of a solid physical object. Thus, perhaps modules in the mind generate the perception of a solid physical object. This, of course, leads to the question of whether solid physical objects exist in the world independently of the perception and cognition of such objects. Now I will not explore that question in any detail here. The only point I wish to make is that the module (or set of modules) responsible for our cognition of folk physics produces the perception of mid-sized physical objects, and hence the *module or modules may also be the origin of the very concept* of a physical object itself. And if we accept current physics, then mid-sized physical objects are mostly empty space. So it is reasonable to ask: Are mid-sized physical objects really 'out there' at all?

If the suggestion that the mind creates the concept of a solid physical object seems rather surprising, consider the perception and thence the concept of hot and cold. The perceptions of hot and cold are very familiar to us, but 'hot' and 'cold' don't exist 'out there' in the world. All that exists 'out there' are the vibrations of molecules. I suggest that the vibration of molecules has only an arbitrary connection to the perceptions of hot and cold. But the contingencies of the evolution of consciousness were such that the perception of hot and cold now represents the vibration of molecules in conscious awareness.

Now, given the arbitrary relation between molecular vibration and the sensation of hot and cold, we might expect other arbitrary relations to exist between our consciousness of the world and what is really 'out there' independent from our experience. However, once a representation, like the experience of hot or cold, is in place, it seems to us, as conscious entities, that things really are hot or cold 'out there'. Similarly, things like solid physical objects appear to be 'out there' but perhaps the concept of a solid physical object is just an heuristic concept provided by a module that makes the flux of experience comprehensible. Consider a similar point made by Quine:

> Physical objects are conceptually imported into the situation as convenient intermediaries – not by definition in terms of experience, but simply as irreducible posits comparable, epistemologically, to the gods of Homer . . . . Both sorts of entities enter our conception only as cultural posits. The myth of physical objects is epistemologically superior to most in that it has proved more efficacious than other myths as a device for working a manageable structure into the flux of experience. (Quine, 1963, p. 44)

And if this is the case for the concept of a physical object, it may also be the case for the concept of an agent. If a folk psychology module (or set of modules) does exist, then I suggest that it may produce the very concept of 'an agent with beliefs and desires' in a similar way to the way in which a visual module produces the Müller–Lyer illusion. To unpack this suggestion a little more, one module might produce the concept of a belief and attribute it to a physical system. Then another module might take the concept of a physical system having a belief as an input together with other inputs (e.g. the concept of a physical system having a desire) and output the concept of an agent with that belief–desire pair. Further, consider the mandatoriness and relative inaccessibility of the workings of a module. I suggest that thinking in terms of an agent with beliefs and desires may be mandatory in the same way that seeing the lines of different lengths is mandatory in the Müller–Lyer illusion. Perhaps we cannot help but think in terms of an agent with beliefs and desires.

Some might say that agents with beliefs and desires are clearly 'out there' in the world. But if the concept of 'an agent with beliefs and desires' is the output of a module or set of modules, and furthermore, if the output of that module, or set, is mandatory and the internal workings of the module are inaccessible to conscious awareness, then the perception that an agent is 'really out there' would be indistinguishable from the output of such a module or set of modules. Thus, I suggest, we must seriously consider the possibility that there are no agents 'out there' in the world, but rather, the concept of 'an agent with beliefs and desires' is produced within a module or set of modules in the mind.

Consider another analogy to illustrate the point: the concept of 'the average family' used by social policy makers to make real world decisions. No one expects to find the average family 'out there' in the world. The average family does not exist in the same way that any particular family exists. I suggest that the concept of 'the average family' is similar to the concept of 'an agent with beliefs and desires'. Just as policy-makers can use the concept of the average family to make real world decisions, a human can use the concept of 'an agent with beliefs and desires' to make real world decisions (either about themselves or others). The social policy-maker does not expect to find the average family 'out there' in the world, and perhaps the human individual should not expect to find 'agents with beliefs and desires' out there in the world either.

If this is correct, then concepts of agents with beliefs and desires can be thought of as *heuristic* concepts. Perhaps the very idea of an agent

with beliefs and desires is a fast and frugal way of making sense of a large and complex set of data, rather than an entity that is 'out there' in the world. Thus, I suggest, concepts that we take for granted (such as the existence of agents) are in fact heuristic concepts that modules in our mind have created to make it easier to process the large information stream that we must process to make our way in the world.

Now I will consider this suggestion in the context of Dennett's work on the physical, design and intentional stances.

4. Dennett's Three Stances

Sellars drew a distinction between two rival images, the manifest and scientific, characterizing the manifest image as 'an "inadequate" but pragmatically useful likeness of a reality which first finds its adequate (in principle) likeness in the scientific image' (1964, p. 57). For Sellars, science is in the process of constructing the scientific image. But the image of the world created by these scientific theories is very different from the common sense or 'folk' image of the world. The common sense or 'folk' image of the world is Sellars's manifest image.

I suggest the distinction between the manifest and scientific images of the world corresponds to the distinction between the images of the world generated by System 1 and System 2 (where System 1 is understood as the modular mind). And I take Dennett to hold a similar position:

> It is no accident that we have the manifest image that we do; our nervous systems were designed to make the distinctions we need swiftly and reliably, to bring under single sensory rubrics the relevant common features of our environment, and to ignore what we can usually get away with ignoring. . . . The undeniable fact is that usually, especially in the dealings that are most important in our daily lives, folk physics works. Thanks to folk physics we stay warm and well fed and avoid collisions, and thanks to folk psychology we cooperate on multiperson projects, learn from each other, and enjoy periods of local peace. (1987, p. 11)

Both Sellars and Dennett appear to endorse the idea that our mind structures our experience of the world. But I take Dennett's position to be particularly amenable to a modular account. (I think of his 'sensory rubrics' as heuristic outputs of modules.) So to Dennett I now turn, and consider his three stances: the physical, design and intentional stances.

I examine Dennett's three stances in two stages. The first stage introduces them as methods by which our limited cognitive systems process the large amount of information we need to process in order to survive. Limited cognitive systems need to use cognitive resources frugally. Thus I characterize Dennett's three stances as fast and frugal heuristics that are produced by modules. This first stage of the examination makes no claims about the existence or non-existence of the entities that these heuristics use as 'content'. The entities may or may not exist, and this is importantly independent of the practical application of these heuristics. We can productively use these heuristics independently from whether the entities referred to by them actually exist (in the very same way that we can use the concept of the 'average family' independently from whether the average family is 'out there' in the world). The second stage addresses the ontological status of the content of the intentional stance (Dennett's 'real patterns') and this will be relevant to the attribution of agency in religious belief.

So I begin with the purely pragmatic use of the three stances as methods of getting by in the world. We are physical organisms and as such we must expend energy to engage in cognition, but we have limited energy so we must be frugal. By contrast, we can conceive of a conscious entity that has unlimited cognitive resources, personified by Laplace's 'demon'. Laplace's demon would see the scientific image of the world. Assuming (incorrectly but for the sake of illustration) a deterministic universe, Laplace's demon could use the scientific image to calculate the future of the universe in all its detail, including the behaviour of humans. But we are not Laplacian demons, so we must employ some cognitive shortcuts, some 'fast and frugal' heuristics. Dennett's three stances are examples of these heuristics.

These stances are practical ways of calculating the behaviour of a range of physical systems, from simple to complex. The stances are practical because without them some physical systems would be too complex for humans to predict. The physical stance is the most basic. This stance assumes all objects of experience are simply physical objects. I take the physical stance to be the output of the so-called 'folk physics' module (or set of modules).

All objects can be considered from the physical stance, however if one's cognitive capacities are limited, it may not be possible for one to understand the behaviour of a complex physical system using the physical stance alone. Dennett illustrates this by contrasting the limited cognitive capacities of a normal human with an omniscient physicist (Laplace's demon) who could predict the behaviour of any physics

system 'assuming it to be ultimately governed by the laws of physics' (Dennett, 1987, p. 23).[1]

If a physical system is too complex to predict using the physical stance, we can move to the design stance 'where one ignores the actual (possibly messy) details of the physical constitution of an object, and, on the assumption that it has a certain design, predicts that it will behave as it is designed to behave' (Dennett, 1987, pp. 16–17). And if the physical system is yet more complex still, even the design stance will not work. Then Dennett moves to the intentional stance. Here is Dennett on taking the intentional stance:

> first you decide to treat the object whose behaviour is to be predicted as a rational agent; then you figure out what beliefs that agent ought to have, given its place in the world and its purpose. Then you figure out what desires it ought to have, on the same considerations, and finally you predict that this rational agent will act to further its goals in the light of its beliefs. (Dennett, 1987, p. 17)

Thus we can see how the three stances can be employed pragmatically to help us, as limited cognitive systems to survive in a complex world.[2]

But, setting aside the utility of these stances, the question remains concerning the reality, or otherwise, of the contents of these stances. For example, the intentional stance attributes beliefs and desires to certain physical systems. But do these systems really have these beliefs and desires? Or are beliefs and desires simply heuristic concepts employed by the modules that generate these stances (in the very same way that a statistician might use the concept of the average family without expecting to find such a family out there in the world)? It is to these questions that I now turn.

5. The Ontological Status of Real Patterns

For Dennett a system 'really' has beliefs and desires if you can successfully take the intentional stance toward that system. But what does 'successfully' mean here? For Dennett, if you are successful in predicting the behaviour of the object (or system) then that object (or system) has beliefs and desires:

> Any system whose behaviour is well predicted by this strategy is in the fullest sense of the word a believer. What it is to be a true believer is to be an intentional system, a system whose behaviour

> is reliably and voluminously predictable via the intentional strategy. (1987, p. 15)

So it is the successful predictions generated by taking the intentional stance toward a system that justifies the system being called an intentional system (a system with beliefs and desires).

As we saw above, Laplace's demon would not need to adopt the intentional stance, simply because it could predict the behaviour of the system from the laws of physics and the initial conditions of the system. But Dennett stresses that if Laplace's demon (now personified as a Martian) failed to take the intentional stance towards the system it was predicting it would be failing to see a 'real pattern' that existed 'objectively' in the world:

> Our imagined Martians might be able to predict the future of the human race by Laplacean methods, but if they did not also see us as intentional systems, they would be missing something perfectly objective: the patterns in human behaviour that are describable from the intentional stance, and only from that stance, and that support generalizations and predictions. (1987, p. 25)
>
> Where there are intelligent beings, the patterns must be there to be described, whether or not we care to see them. It is important to recognize the objective reality of the intentional patterns discernible in the activities of intelligent creatures . . . (p. 28)
>
> I claim that the intentional stance provides a vantage point for discerning . . . useful patterns. These patterns are objective – they are *there* to be detected – but from our point of view they are not *out there* entirely independent of us, since they are patterns composed partly of our own 'subjective' reactions to what is out there . . . It is easy for us, constituted as we are, to perceive the patterns that are visible from the intentional stance – and only from that stance. (p. 39)

Dennett concedes that these patterns are not 'out there' entirely independent of us. So what is the nature of this objectivity? Dennett describes these 'real patterns' as partly composed of our own 'subjective' reactions to what is out there. This I think is an interesting approach. Dennett is claiming that something about the nature of the human condition (our ability to adopt the intentional stance) is central to the reality of these patterns. Now I want to push Dennett's position slightly to see where we can take it. In particular, I want to explore the

possibility of granting this semi-real status to spirits associated with natural systems that are common in religious belief. I am claiming that something about the nature of the human condition (our ability to adopt the spiritual stance) is central to the reality of gods and spirits. And here we move toward the second of the questions at the heart of this chapter: Why do we attribute religious agency?

6. Changing the Measure of Success

Dennett justifies the attribution of beliefs and desires to a system if we can accurately predict the behaviour of the system using those attributions. But why is predictive success the appropriate justification? The ability to predict successfully their environment is important to organisms making their way in the world. But it is certainly not the only important ability. The most important ability is reproductive success. Indeed, it could be argued that the ability to predict successfully systems in their environment is only important if it furthers the reproductive success of the organism.

So, moving from Dennett's instrumentalism to my pragmatism, I change the measure of success from predictive success to reproductive success over evolutionary time. Now with this new measure of success in place, a system has beliefs and desires (still in Dennett's semi-realist sense) if attributing beliefs and desires to that system leads to reproductive success for the attributing population of organisms over evolutionary time.

Now critics (perhaps Dennett included) might say this has little to do with Dennett's original intentional stance. Indeed, Dennett warns of radical interpretationalism (1987, p. 24) of which this may be a species. But in response to such concerns I simply point out that, from an evolutionary point of view, reproductive success is all that matters. Predictive success may lead to reproductive success, but at the end of the day it is reproductive success that determines the long-term existence of any lineage of organisms. And, as is pointed out by Wilson and Lynn, this has implications for what we take to be *real*:

> Dozens of evolutionists have observed that insofar as beliefs are products of natural selection, either proximally or distally, then they should be designed to enhance fitness, not to perceive the world as it really is. . . . To begin with genetic evolution . . . deception begins with perception. All organisms perceive only the environmental stimuli that matter to their fitness. Our species can see only a narrow slice

> of the sound and light spectrum, cannot sense electrical and gravitational fields at all, and so on. We also distort what we can perceive, for example, by turning the continuous light spectrum into discrete colors. Perception might not qualify as belief, but if the former is so prone to adaptive distortions, it would be surprising if the latter was not prone as well. (Wilson and Lynn, 2009, p. 539)

Of course, Dennett may wish to distinguish 'real patterns' that are realized by predictive success from 'real patterns' that are realized by reproductive success, and it is reasonable to do so. So, to distinguish this stance (with its different measure of success) I call it the spiritual stance.

Importantly the only thing I have changed in comparison to the intentional stance is the measure of success. Everything else remains as Dennett characterized it. So the spiritual stance attributes beliefs and desires to systems and if that attribution leads to reproductive success over evolutionary time, then the system in question has beliefs and desires in the semi-realist sense that Dennett describes. So a successful application of the spiritual stance leads to the existence of real spiritual patterns.

7. Why Would the Spiritual Stance Lead to Reproductive Success?

The spiritual stance is a way of understanding the attribution of agency in religious contexts. It is also a way of understanding the (semi-real) existence of gods and spirits as real patterns. But all this presupposes that the attribution of religious agency would lead to reproductive success over evolutionary time. So, let me now turn to the question, Would the spiritual stance lead to reproductive success?

A position gaining acceptance within cognitive science of religion makes the following claims: (1) religious belief supports cooperative behaviour in groups (Johnson and Bering, 2009); and (2) cooperative groups outcompete non-cooperative groups (via group selection); and thus (3) religious belief was selected over evolutionary time (Wilson, 2002). So, it is claimed, religious belief has led to reproductive success. This position is reasonably well established in the cognitive science of religion literature, so I will not examine it in any detail here, but see Wood (2011) for a review. However, I will say a few words about how the spiritual stance fits into this account of religious belief.

The position outlined above focuses on belief in the existence of punishing gods and their role of in the facilitation of in-group cooperation.

At first sight there does not seem to be much in common with that position and the spiritual stance, in which beliefs and desires (and agency) are attributed to physical systems. However, with only a small change in the content of the relevant belief system one can get from an account involving a physical system to an account involving punishing gods. For example, if the physical system toward which the spiritual stance is taken is, say, the local physical environment (or alternatively, the universe as a whole), and the beliefs and desires attributed to it relate to a belief about an individual cheating in a group task, and the desire to punish the cheat, then it is only a small change in the relevant belief system to get from the local physical environment to a local god, as might be relevant to an ancient hunter-gather culture (or alternatively, from the universe as a whole to a god 'outside' the universe, as might be relevant to classical monotheism).

However, there are many religious beliefs that do not involve gods and spirits that desire to punish, so now I suggest how the spiritual stance might lead to reproductive success without relying on the notion of gods that punish. The suggestion involves a particular concept of personhood, namely the concept of *an agent with moral standing*. And I suggest that attributing personhood (in this particular sense) to certain natural systems may lead to reproductive success over evolutionary time.

To clarify this concept of personhood consider two distinctions. The first is between humans and agents. I take humans to be defined *biologically* with reference to the fact that they are members of the species *Homo sapiens*. I take an agent to be defined *intentionally* with reference to the fact that the intentional or spiritual stance is successfully taken toward them. Thus a human is not essentially an agent, and an agent is not necessarily a human.

The second distinction is between agents and persons. Again while agents are understood intentionally, persons are understood morally. I take personhood essentially to involve moral standing being bestowed upon an agent (human or otherwise). Thus an agent is not essentially a person. Rather personhood is bestowed upon an agent (human or otherwise), either by the agent themselves or other agents. Alternatively, agents may choose not to bestow personhood on themselves or other agents (or other systems). And importantly, personhood can be bestowed on any physical system (assuming agency has also been attributed to it).

So far, I have kept the concept of attributing agency and attributing personhood distinct. It is conceptually possible to attribute agency without attributing personhood. But I suggest that there is a deep connection between attributing agency and attributing moral standing to

that system. I suggest it is *natural* to attribute moral standing to a system when agency is attributed to that system.

I think that attributing personhood to a system is one of the most basic cognitive attitudes that we can take towards a system with which we are interacting. I also suggest that the process of attributing personhood to a system is closely related to the modular process described above whereby a module attributes beliefs and desires to a physical system. Just as System 1 may be responsible for attributing agency to a physical system, it may also be responsible for attributing moral standing to that system, and then representing that system as a person (an agent with moral standing) to System 2.

Now consider the following question: Would the attribution of personhood to a particular natural system lead to the reproductive success of the linage of humans that made that attribution? In certain circumstances, I suggest that it would, and here is why. If one attributes personhood to a system then one may act differently toward that system. For example, consider a population of humans that grant personhood (and hence moral standing) to the rivers or forests with which they interact. This may change the way they interact with those systems. They might, for example, refrain from using all the resources available within that system. By doing so they may interact sustainably with the system and thus these organisms may survive over evolutionary time.

Furthermore, imagine a number of different populations of organism that attributed personhood to different systems. If a particular attribution of personhood to a system led to sustainable resource use, then the population that made that attribution would persist over evolutionary time, while a different population that did not make that attribution may deplete their resources (even if they flourished for a short period while they depleted them) and thus they may tend toward extinction. Over evolutionary time populations that have a propensity to make attributions of personhood that lead to reproductive success may displace populations that have no such propensity. And thus populations that afford moral status to physical systems upon which their existence depends may flourish in the long term.

8. Conclusion

I have offered a characterization of the attribution of religious agency. I have drawn from the insights of evolutionary psychology and philosophy of mind. Evolutionary psychology suggests that rather than think of the mind as a single cognitive system, we should think of the mind

as many distinct systems that evolved in response to distinct survival challenges. These distinct systems can be conceptualized in a number of ways. One way is the dual process theory of mind where System 1 generates fast and frugal intuitions and System 2 generates all things considered judgements. Another way involves the existence of modules in the mind. Furthermore, the operation of such modules may be mandatory and conscious thought may have limited access to the modular processes. I characterized System 1 as generating jigsaw pieces that were given to System 2. But System 2 does not necessarily understand that System 1 itself constructed the jigsaw pieces. System 2 (the conscious mind) then puts those pieces together to make sense of its world.

In addressing the question 'How do we attribute agency?' I introduced Dennett's intentional stance – the attribution of beliefs and desires (and hence agency) to physical systems – and suggested that this stance was produced by a set of modules in the mind. Due to the mandatory function of the modules and the fact that conscious awareness has limited access to the functioning of these modules, conscious awareness only sees the output of this modular system and thus takes humans to have belief, and desires (and agency). I also explored the ontological status of beliefs and desires as 'real patterns'.

Then moving toward the question 'Why do we attribute agency?' I adapted the intentional stance by changing the measure of success from 'predictive success' to 'reproductive success over evolutionary time' and thereby I introduced the spiritual stance. But the spiritual stance preserved all of the implications of the original intentional stance. Thus, if the attribution of beliefs and desires (and hence agency) to a physical system led to reproductive success over evolutionary time, then that physical system had beliefs and desires (and hence agency) in Dennett's semi-realist sense. Drawing on insights from the cognitive science of religion, and using a concept of personhood that bestowed moral standing on physical systems attributed with agency, I suggested that there could be circumstances in which such personhood/agent attributions did lead to reproductive success over evolutionary time. In conclusion, all this is a way to understand both the attribution of agency in religious contexts and the resultant (semi-real) existence of gods and spirits as 'real spiritual patterns'.[3]

Notes

1. The full passage reads as follows: 'The Laplacean omniscient physicist could predict the behaviour of a computer – or of a live human body, assuming it to be ultimately governed by the laws of physics – without any need for

the risky, short-cut methods of either the design or intentional strategies' (Dennett, 1987, p. 23). Now, this passage reads as if the physical stance is read directly off the 'laws of physics' and on this point my analysis differs from Dennett's. I take the physical stance to be distinct from the scientific image. On my account, the physical stance is the output of the 'folk physics' module. But it seems that Dennett takes the physical stance to be read directly off the scientific image. But I have set this difference aside for the purposes of this chapter.

2. There is some difference between Dennett's account and my account here. Dennett's account implies that we can choose to swap stances, but it is not clear to me whether the content of the three stances are produced by central cognition or by modules (what Dennett called the manifest image). On my account the core content (what I have called jigsaw pieces) of all three stances are produced by modules, so the content is given to central cognition (cf. the 'given' in Sellars). But central cognition might choose to ignore the output of a particular module or combine the outputs of modules in various ways (what I have called putting the jigsaw together).
3. This chapter is an output from a project undertaken as part of the Cognition, Religion and Theology Project at the University of Oxford, funded by the John Templeton Foundation. The views expressed are not necessarily those of the Cognition, Religion and Theology Project, the University of Oxford or the John Templeton Foundation.

References

Baron-Cohen, S. (1995) *Mindblindness: An Essay on Autism and the Theory of Mind* (Cambridge, MA: MIT Press).

Carruthers, P. (2006) *The Architecture of the Mind* (Oxford: Clarendon Press).

Cosmides, L., and J. Tooby, 'Foreword', in S. Baron-Cohen (1997) *Mindblindness: An Essay on Autism and the Theory of Mind* (Cambridge, MA: MIT Press), pp. xi–xviii.

Dennett, D. (1987) *The Intentional Stance* (Cambridge, MA: MIT Press).

Fodor, J. (1983) *The Modularity of Mind* (Cambridge, MA: MIT Press).

Gerrans, P. (2002) 'The Theory of Mind Module in Evolutionary Psychology', *Biology and Philosophy* 17: 305–21.

Gigerenzer, G., P. M. Todd et al. (1999) *Simple Heuristics That Make Us Smart* (New York: Oxford University Press).

Gilovich, T., D. Griffin and D. Kahneman (2002) *Heuristics and Biases: The Psychology of Intuitive Judgment* (Cambridge: Cambridge University Press).

Hookway, C. (2010) 'Pragmatism', in E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy.* Available at http://plato.stanford.edu/archives/spr2010/entries/pragmatism/ (accessed 20 October 2011).

James, W. (1975) *Pragmatism and the Meaning of Truth* (Cambridge, MA: Harvard University Press).

Johnson, D., and J. Bering (2009) 'Hand of God, Mind of Man: Punishment and Cognition in the Evolution of Cooperation', in J. Schloss and M. Murray (eds), *The Believing Primate* (Oxford: Oxford University Press), pp. 26–43.

Kahneman, D. (2002) 'Maps of Bounded Rationality: A Perspective on Intuitive Judgement and Choice', *Nobel Prize Lecture*. Available at http://nobelprize.org/nobel_prizes/economics/laureates/2002/ kahnemann-lecture.pdf (accessed 28 March 2011).

Quine, W. V. (1963) *From a Logical Point of View* (New York: Harper & Row).

Richardson, R. (1999) 'Heuristic', in R. Audi (ed.), *The Cambridge Dictionary of Philosophy* (Cambridge: Cambridge University Press, 1999), p. 379.

Samuels, R. (2000) 'Massively Modular Minds: Evolutionary Psychology and Cognitive Architecture', in P. Carruthers and A. Chamberlain (eds), *Evolution and the Human Mind: Modularity, Language and Meta-Cognition* (Cambridge: Cambridge University Press), pp. 13–46.

Samuels, R., S. Stich and P. Tremoulet (1999) 'Rethinking Rationality: From Bleak Implications to Darwinian Modules', in E. Lepore and Z. Pylyshyn (eds), *What Is Cognitive Science?* (Oxford: Blackwell), pp. 74–120.

Segall, M., D. Campbell and M. J. Herskovits (1966) *The Influence of Culture on Visual Perception* (New York: The Bobbs-Merill Company).

Sellars, W. (1964) 'Philosophy and the Scientific Image of Man', in R. Colodny (ed.), *Frontiers of Science and Philosophy* (London: Allen & Unwin), pp. 35–78.

Simon, H. (1962) 'The Architecture of Complexity', *Proceedings of the American Philosophical Society* 106: 467–82.

Spelke, E. S. (1990) 'Principles of Object Perception', *Cognitive Science* 14: 29–56.

Stanovich, K., and R. West (2000) 'Individual Differences in Reasoning: Implications for the Rationality Debate', *Behavioural and Brain Sciences* 23: 645–65.

Wilson, D. S. (2002) *Darwin's Cathedral* (Chicago: University of Chicago Press).

Wilson, D. S., and S. J. Lynn (2009) 'Adaptive Misbeliefs Are Pervasive, but the Case for Positive Illusion Is Weak', *Behavioural and Brain Sciences* 32: 539–40.

Wood, G. (2011) 'Cognitive Science and Religious Belief', *Philosophy Compass* 6.10: 734–45.

This page intentionally left blank

Part III
God and the Universe

This page intentionally left blank

5
On Non-Singular Space-times and the Beginning of the Universe

William Lane Craig and James D. Sinclair

1. Introduction

The most widely discussed argument for the existence of God is the so-called *kalam* cosmological argument, which originated in attempts on the part of certain ancient philosophers to rebut Aristotle's doctrine of the past eternity of the universe.[1] The argument assumed major importance in medieval Islamic theology, from which its name derives. In his *Kitab al-Iqtisad* the medieval Muslim theologian al-Ghazali presented the following simple syllogism in support of the existence of a Creator: 'Every being which begins has a cause for its beginning; now the world is a being which begins; therefore, it possesses a cause for its beginning' (1962, pp. 15–16). In defence of the second premise, Ghazali offered various philosophical arguments to show the impossibility of an infinite regress of temporal phenomena and, hence, of an infinite past. The limit at which the finite past terminates Ghazali calls 'the Eternal' (1963, p. 32), which he evidently takes to be a state of timelessness. Given the truth of the first premise, the finite past must, therefore, 'stop at an eternal being from which the first temporal being should have originated' (1963, p. 33).

The argument, then, is disarmingly simple:

1. Everything that begins to exist has a cause.
2. The universe began to exist.
3. Therefore, the universe has a cause.

Conceptual analysis of what it means to be a cause of the universe then seeks to explore the relevance of this conclusion for theism. What

makes the *kalam* cosmological argument scientifically interesting is that contemporary proponents of the argument have claimed support from current astrophysical cosmogony in defence of premise 2. A number of recent critics of the argument, however, have denied that contemporary physical cosmology lends any significant support to premise 2 because the singular origin of space-time predicted by the standard model is an artefact of General Relativity (hereafter GTR) which may well be resolved by quantum cosmological models. For example, Bradley Monton states,

> The big bang hypothesis is true given the assumption that GTR is true, but we don't know that the big bang hypothesis is true of the actual universe. The big bang theory doesn't take into account quantum theory, and that gives us reason not to believe the big bang theory.... Because the physics doesn't tell us what happens once we trace the history of the universe backwards in time to these high energies, we don't even know if there's a big bang at all. (forthcoming)

Monton appeals specifically to the Steinhardt–Turok cyclic ekpyrotic model as an example of a cosmological model which 'is compatible with the universe having been in existence forever' (forthcoming). Similarly, Graham Oppy complains that we are 'far from having good reason' to suppose that quantum gravitational replacements of the standard model will feature an absolute beginning of the physical universe' (2006, p. 146). Undoubtedly, the most extensively developed criticism of this sort comes from J. Brian Pitts in his wide-ranging critique of what he styles 'the singularity argument for theism' (2008, p. 677). His critique therefore merits careful scrutiny.

It is a curious feature of Pitts's lengthy critique that he never actually provides an explicit statement of the target singularity argument for theism. At first blush, one might think that the *kalam* cosmological argument is just the singularity argument for theism. But a moment's reflection shows that this cannot be right, for, at most, evidence for an initial cosmological singularity might be adduced in support of premise 2. It is immediately obvious, then, that 'the singularity argument for theism' is a misnomer, for the singularity argument is not an argument *for theism*. Rather it is an argument for premise 2, which is a straightforward, physical statement, the likes of which is discussed in any text on astronomy and astrophysics and which does not posit any supernatural entity. What is really at issue, then, is some sort of singularity argument for the beginning of the universe.[2] The recognition of this

fact is highly significant, since it implies that all of Pitts's worries about 'God-of-the-gaps' reasoning and the impenetrability of natural theology (2008, pp. 696–700) are simply irrelevant to the success of the so-called singularity argument for the beginning of the universe and so may be left aside. In what follows, then, we shall have very little to say of theology. This is, we think, all for the better, for the question of whether contemporary cosmology furnishes good grounds for thinking that the universe began to exist is of more than sectarian interest.

2. The Singularity Argument Formulated

Our concern, then, is with some supposed singularity argument for the beginning of the universe. In order to discern the alleged failings of that argument, we need a clear formulation of the argument. This Pitts unfortunately does not provide, leaving us to reconstruct it as best we can. The key to reconstructing the so-called singularity argument appears to be Pitts's summarizing statement:

> Whether one is tolerant or intolerant toward singularities, it turns out that there is no first moment (unless one is installed by hand), because every moment is preceded by earlier moments. Thus in the relevant sense for the Kalam argument to be valid and to make sense for a sufficiently broad collection of physical theories, there is no beginning implied by physics, and so premise 2 might be false, as far as physics can show. In order for the Big Bang singularity to provide a good theistic argument, the singularity must be well enough behaved to be a real and intelligible part of space-time and badly enough behaved that it cannot have a past. Satisfying both conditions seems difficult and unlikely to be achieved. (2008, p. 689)

This passage might suggest the following reconstruction:

4. The universe began to exist if there was a first moment of time.
5. If the initial cosmological singularity was a real and intelligible part of space-time, then there was a first moment of time.
6. The initial cosmological singularity was a real and intelligible part of space-time.
7. Therefore, the universe began to exist.

The problem with this suggested reconstruction of the argument, however, is that the argument would then be a straw man which has been defended by no one and would hardly require 21 pages of sustained

argumentation to refute. Singularities of general relativistic space-times are not space-time points but, at most, points, whether real or ideal, which can be attached to space-time as boundary points. As such, they are not parts of space-time, as (6) asserts. Indeed, on the customary approach to singularities, they are not even accorded that much reality. Rather a space-time is said to be singular if it is essentially inextendible and so geodesically incomplete (Earman, 1995, pp. 28–37). The question Pitts raises is whether an initial cosmological singularity, either in the sense of a past boundary point or of the incompleteness of past geodesic half-curves, gives grounds for thinking that the universe began to exist. Asking the question in this way renders intelligible Pitts's further worry whether some as yet unknown theory might not resolve the singularities of classical space-time – a worry which would become superfluous with respect to the above proposed reconstruction – for even if an initial cosmological singularity implies the beginning of the universe, the question remains whether space-time really is singular.

Accordingly, we might offer the following reconstruction of the supposed singularity argument for the beginning of the universe:

8. If space-time is singular, the universe began to exist.
9. Space-time is singular.
10. Therefore, the universe began to exist.

Pitts may then be understood to deny (8) on the grounds that, whether the initial singularity is understood as a boundary point to the past or as the past geodesic incompleteness of space-time, there is no first moment of the universe's existence, which he takes to be a necessary condition of the universe's beginning to exist, and to undercut (9) by appealing to as yet undiscovered theories which remove the singularity. This seems to us a sympathetic reconstruction of the alleged singularity argument for the beginning of the universe which also makes good sense of Pitts's two-pronged response (Pitts, 2008, p. 682).

3. The Singularity Argument Assessed

Even so, it must be said that the singularity argument for the beginning of the universe is still a straw man. For the evidence that the universe began to exist does not depend on evidence for a *singular* beginning of space-time. We have elsewhere discussed a variety of non-singular cosmogonic models and have noted the difficulties in plausibly tracing out an infinite past (Copan and Craig, 2004; Craig, 2006; Craig and Sinclair,

2009). We shall have more to say of this in the sequel, when examining what support contemporary physical cosmology lends to premise 2.

3.1. Assessment of premise 8

Taking the argument on its own terms, however, let us first examine Pitts's objection to (8). Despite his flirtation with metric conventionalism, Pitts recognizes that our universe is characterized by an objective temporal metric. A Friedman–Robertson–Walker universe, being bounded by an initial cosmological singularity, is metrically finite in the past. Following Smith (1985), we can say plausibly that time begins to exist if for any arbitrarily designated, non-zero, finite interval of time, there are only a finite number of isochronous intervals earlier than it; or, alternatively, time begins to exist if for some non-zero, finite temporal interval there is no isochronous interval earlier than it. That condition is satisfied in the standard model of space-time. So why is the metrical finitude of the past not sufficient for time's beginning to exist? Pitts's answer is that 'Within the Robertson–Walker cosmological space-time for $t > 0$ (which is to say, always), one can explain each moment in terms of an earlier one. Thus there is no beginning required and premise 2 might be false, as far as physics can tell' (2008, p. 688). Pitts's point is not that each instantaneous slice of space-time can be immanently explained in terms of a prior slice, thereby obviating the need for a transcendent cause of the universe's beginning to exist;[3] rather his point is that since $t = 0$ is at most a boundary point of space-time, there is no first instant of time and thus no beginning of time, even if past time is finite, and, hence, no beginning of the universe.

Pitts's objection presupposes that *beginning to exist* entails *having a beginning point.* But why should we think that? Pitts offers three reasons why construing the universe's beginning to exist in topological rather than metrical terms is theologically advantageous for the biblical doctrine of creation (2008, pp. 680–1). But such a concern is foreign, not only to natural theology, but to the religiously neutral question of whether the universe began to exist. Relevant philosophical considerations are those that have purchase with philosophers and cosmologists regardless of their religious persuasion. What is advantageous to the biblical doctrine of creation does not and should not enter into the question of what it is to begin to exist.[4] Indeed, given the above intuitively plausible sufficient conditions for time's beginning to exist, Pitts's scenario of a metrically finite past lacking a first instant may be taken to provide good reason to conclude that *beginning to exist* does not entail *having a beginning point*.[5]

Moreover, Pitts's entailment would commit us to the reality of points. But whether space and time really are composed of an actual, nondenumerable infinity of points, rather than simply modelled as such in GTR, is surely a question to be settled by argument and evidence, not merely assumed, as Pitts does (2008, p. 680).

Furthermore, Pitts's own Cosmic Destroyer argument (2008, pp. 692–3) seems to undermine his presumed entailment. For if *beginning to exist* entails *having a beginning point*, then, *pari passu*, *ceasing to exist* entails *having an ending point*. Since a black hole singularity (or a terminal cosmological singularity) is at most a boundary point of space-time rather than a part of space-time, any object which falls into a black hole has no ending point of its existence. Nevertheless, it does cease to exist. As Robert Geroch explains,

> Recall there is no event available to [an observer headed toward a singularity] on the singularity itself. In particular, there is no possibility for a further extension of [the observer's] world-line after it hits the singularity. What, then, does 'the world-line of [an observer] hits the singularity' mean physically? Mathematically, what happens is that this world-line just stops. Physically, this would mean that [the observer] is 'snuffed out of existence'; after some finite time according to himself, he ceases to exist in space-time. (Geroch, 1978, p. 194)

This isn't simply rhetorical gloss on Geroch's part. The observer would indeed be 'snuffed out of existence' despite the lack of a genuine ending point to the space-time. If, then, the universe ceases to exist in the Big Crunch, parity requires that it began to exist in the Big Bang, QED.[6]

Finally, it deserves to be pointed out that Pitts's insistence that the crucial consideration as to whether the universe began to exist is the necessity of a first instant lands him squarely in the ancient Greek paradoxes of stopping and starting (see Sorabji, 1983, pp. 403–21). Ancient thinkers like Parmenides argued that if an object O is at rest at time t, it is impossible for O to begin to move, since for any time $t' > t$, if O is in motion at t', then there is a time $t < t^* < t'$ at which O is already in motion. Hence, nothing can ever begin to move. If we assume the continuity of time and space, the solution to Parmenides' puzzle is that O can begin to move without there being a beginning point of its motion. Pitts's demand for topological closure would force upon us the absurd conclusion that nothing ever begins to move. But if we allow that *beginning to move* does not entail *having a beginning point* of motion, then, generalizing, neither should we demand that

in beginning to exist the universe must have a beginning point of its existence.

We need, then, some compelling reason to think that the admitted finitude of past time is not sufficient for time's having begun to exist, for thinking that a first instant of time is a necessary condition as well. Pitts's relevant remarks in this regard come in response to Copan and Craig's definitions which they provide in preparation for their defence of premise 3* (the temporal series of past physical events was not beginningless). They explain:

> By a 'physical event,' we mean any change occurring within the space-time universe. Since any change takes time, there are no instantaneous events. Neither could there be an infinitely slow event, since such an 'event' would in reality be a changeless state. Therefore, any event will have a finite, nonzero duration. In order that all the events be of equal duration, we arbitrarily stipulate some event as our standard.
>
> Taking as our point of departure the present standard event, we consider any series of such standard events ordered according to the relation *earlier than*. The question is whether this series of events had a beginning or not. By a 'beginning' one means a first standard event. It is therefore not relevant whether the temporal series had a beginning point (a first temporal instant). The question is whether there was in the past an event occupying a nonzero, finite temporal interval that was absolutely first, not preceded by any equal interval. (Copan and Craig, 2004, p. 199; emphasis original)

Pitts's objection to this *Ansatz* is surprising. He writes,

> The Bach–Weyl theory of gravity shows that a physical theory need not even define the length of a curve. In scalar-tensor theories or other theories with multiple metrics, age might be radically ambiguous. Thus, only a topological notion of 'beginning' in terms of a first moment is available if the Kalam argument is intended to yield a necessarily nonvacuous necessary truth ....
>
> Copan and Craig perhaps take their Kalam argument to express a necessary truth that applies nonvacuously to all possible physical theories. However, their criterion for a beginning is meaningless for a Bach–Weyl theory because 'equal duration' is meaningless in that theory .... Given that neither existence nor uniqueness of a metric (for timelike curves) holds necessarily, the natural move is to adopt a topological rather than metrical notion of beginning. Thus a first

> moment *is* the point that needs to be addressed. (2008, pp. 677, 682; emphasis original)

Pitts's presumption seems to be that Copan and Craig take not only the premises of the *kalam* cosmological argument, but more particularly their proffered criterion for there being a beginning of the series of past events to be metaphysically necessary truths. Since there are possible worlds characterized by metric conventionalism – that is to say, worlds in which there is no objective fact of the matter whether or not non-nested temporal intervals are isochronous – the proffered criterion for a beginning would be inapplicable to such worlds, for there just is no fact of the matter in such worlds whether any other temporal interval is equal to the duration of the arbitrarily stipulated, standard event. Pitts recognizes, as mentioned, that the actual world is characterized by an objective metric; but he insists that 'a doctrine of creation needs to be modally rich enough to accommodate the possibility of God's creating worlds with physical laws without a unique or preferred metrical structure to license an answer of 'finite' or 'infinite' age' (Pitts, 2008 p. 681).

Pitts's presumption is wrong. The proffered criterion was intended only to apply to the actual world. There is no intention or need to seek its application to other metaphysically possible words governed by different laws of nature.[7] Thus, the proffered criterion for a beginning of the series of past events is adequate to its purpose. It states a sufficient condition for time's having begun to exist.

In sum, Pitts's attempt to discredit an intuitively plausible sufficient condition of time and the universe's having begun to exist in terms of metrical past finitude rests upon an ungrounded presumption about the modal status of the proffered condition and would substitute a topological condition which, if accepted, would have highly implausible consequences. Hence, he presents no serious objection to (8).

3.2. Assessment of premise 9

What, then, about (9)? If (9) is true, then, as we have seen above, we do have good reason to infer that (10) is true, that the universe began to exist, which is the second premise of the *kalam* cosmological argument. Unfortunately, Pitts seems to assume that the evidential support for premise 2 depends crucially on the cogency of his singularity argument, so that if the warrant for (9) is undercut, so is the warrant for (2). But this assumption is mistaken. Premise 2 does not presuppose or require the truth of (9) or the cogency of the singularity argument. Evidence of a non-singular beginning would be quite sufficient to warrant belief in (2).

Of course, if (9) is unwarranted, then we must ask what support physical cosmology lends to (2). Pitts is worried by the fact that

> As one sees all the time in papers on quantum gravity, most people who work on quantum gravity take for granted that the Big Bang singularity is an artifact of incomplete physical understanding and expect or hope that uniting gravity with quantum mechanics in some kind of quantum gravity will resolve the singularity into some well-defined situation that admits to still earlier times, *ad infinitum*. (2008, p. 688)[8]

The crucial consideration here will not be Pitts's worry that a successful theory of quantum gravity will resolve the initial cosmological singularity featured in the standard model into a situation which is physically well defined, but rather whether, as he puts it, that situation 'admits extrapolation to still earlier times, *ad infinitum*'. Since it is that consideration, rather than the truth of (9), that is crucial to the evidence for (2), let us turn to an examination of that consideration. Do models featuring an era 'before the Big Bang' imply a beginningless past?

4. Physical Evidence for Premise 2

We shall now consider some contemporary cosmogonies.[9] The material in sections 4.2–6 is not intended as an inductive attempt to show that all contemporary cosmogonies may lend support to (2). Rather, it is intended to support the view evinced in Craig and Sinclair, 2009, that there are general principles regarding cosmogonies that yield a taxonomy of models based on the models' expansion behaviour which speaks to the question of temporal origin. We will hold that there is a genuine unique objective past (temporally ordered), with evolutionary continuity to the present. This preserves our veridical experience of time[10] and permits theories of evolution (including Darwin) elsewhere in physics to be understood on a realist footing, as opposed to some type of illusion. Thus, for example, if a particular theory denies temporality as a 'fundamental' feature of the universe, but preserves the notion as a quasi-classical 'approximation', we will take the quasi-classical history as real and suggest the more general formalism should be interpreted instrumentally.

Further, as Pitts seems to assert without evidence that a pre-Big Bang timeline would, of necessity, be past eternal, we shall content ourselves

to providing a few counterexamples as opposed to attempting to prove that *all* pre-Big Bang models fail to demonstrate past eternality.

4.1. Models, methodology, and their use

In the sections following this one, we shall consider the evidence in favour of the proposition that the universe has a temporal origin. Since the consideration of such evidence commands the bulk of our essay, we have little space to defend our implicit presupposition of time's reality. One might base an 'atemporalist' position on the idea that a successful theory of quantum gravity will cease to regard time as a fundamental characteristic of nature (Kiefer, 2008). Here we offer the following brief comments on the atemporalist position:

1. *The atemporalist position is uncritically literalist with regard to model interpretation*. Our interest is in which cosmogonic model most nearly approximates reality. The mere existence of a useful mathematical formalism does not imply that the entities in that theory correspond with reality. The past (and present) usefulness of GTR, for example, never meant that we were obligated to make an ontological commitment to its view that gravity *just is* the curvature of an objectively real space-time. Nor are we obligated to believe that time does not exist because a 'pure' Wheeler–DeWitt formalism does not contain it and the WKB surrogate is described as merely an 'approximation'. The Wheeler–DeWitt quantum gravity approach describes both the disappearance of absolute time that results from the attempt to quantize Einstein's gravity as well as time's reappearance as a semi-classical approximation. The fact that time is recoverable as an approximation, however, does not imply that there is either an elimination of time ('pure' Wheeler–DeWitt) or that the superposition of all possible metrics necessarily requires an indeterminate meaning to time *in reality*. Useful models exist at different levels of abstraction from physical reality. We therefore feel perfectly comfortable invoking the approach of Kiefer (and his sometime colleague H. D. Zeh) in our discussion of cyclic universes and a reversed arrow of time (see section 4.6) on the grounds that one need not attach ontological commitment to their larger metaphysical claims regarding the elimination of time.

   The issue at hand is therefore not the narrow issue of whether there are quantum gravity modelling schemes that are atemporalist. Rather the question is what determines our ontological commitments. Although contemporary philosophy of time has been deeply

imbued with the naturalized epistemology of W. V. Quine (1969), which shuns so-called 'first philosophy' in favour of taking the deliverances of the natural sciences as authoritative guides to reality, such epistemological naturalism is by its very nature simply a methodological disposition, which cannot itself be grounded either scientifically or in any other way, to restrict one's basic sources of evidence to the natural sciences (Rea, 2002, pp. 63–7; cf. pp. 1–7). Not being so disposed, we see no reason that physics should subvert the veridicality of our experience of tense and temporal becoming, much less of time itself. Indeed, even the illusion of temporal becoming entails the reality of temporal becoming in the contents of consciousness. We find it odd that it is 'ordinary' quantum mechanics which brings to the table a near Newtonian view of absolute time, yet this is the ingredient that supposedly spoils the notion of temporality. It is 'ordinary' quantum mechanics that enjoys the impressive experimental verification. We should take care in transferring a borrowed legitimacy to a quantum gravity theory that itself lacks consensus.

It seems to us appropriate to consider each and every scheme that cosmologists bring to the table to throw light on the question of which models approximate physical reality. We are not wedded to the specific concept of metric time found in GTR. Authors of cosmogonies routinely attempt to address either the origins question itself or, at least, questions regarding the universe's initial conditions. Many quantum cosmogonies include claims of past eternality (and thus presuppose the reality of time). Others discuss cosmic beginnings *ex nihilo*. Surely the cosmologists who formulate these theories think that they are doing work relevant to the actual universe and not merely mathematical exercises that have no basis in reality. Hence, it seems to us a profitable and important exercise to address such models.

2. *Quantum Gravity models must replicate the observed behaviour that more primitive models of the universe (read GTR) seek to model*. Isaac Asimov, in an article entitled 'The Relativity of Wrong' (Asimov, 1989), pointed out the fallacy of believing that current 'primitive' models lack usefulness because they are 'wrong', given that a future theory is always just over the horizon to overturn the current one. There are degrees of 'wrongness' and theories which are less wrong than their predecessors are so because they take account of what is empirically correct in their forebears. It is simply mistaken to think that phenomena predicted and analysed by GTR (especially indirectly *observable* phenomena such as black holes) via such things as singularity

theorems cease to be relevant because a different mathematical formalism may someday come into general use. There may be no such things as singularities per se in a future quantum gravity formalism, but the phenomena that GTR incompletely strives to describe must nonetheless be handled by the refined formalism, if that formalism has the ambition of describing our universe. This can be seen in the above example of the Wheeler–Dewitt approach as described by Kiefer. Big Bang Friedmann universes with a meaningful time coordinate are recoverable in the theory (as Kiefer shows) but are described as a semi-classical WKB approximation which invokes a mechanism called decoherence. They in fact must be recoverable because that is what we *see*. If Big Bang Friedmann-like behaviour were not recoverable, that in itself would be grounds for rejecting the larger theory.

3. *Preservation of time as a fundamental property of the universe is possible within a quantum gravity approach* (as examples, see Mersini-Houghton, 2009; Smolin, 2009; Sorkin and Rideout, 1999; Carroll, 2008; Loll, 2008). Mersini-Houghton defends an arrow-less but nonetheless global and beginningless temporal axis within a larger multiverse. Carroll wants a global time with an arrow that points toward greater entropy given any arbitrary point of departure within the multiverse. Smolin defends a similar view in his postulation of multiple universes birthed in a temporal sequence through black hole incubators. His article 'The Unique Universe' is a full-fledged attempt at refutation of a timeless multiverse. Yet his defence of the concept of fundamental time is *not* a defence of GTR's metric time. Rather it defends temporality within the context of *background independent* (i.e. emergent space-time) quantum gravity theories. An example of this approach is the work of Renate Loll which we briefly discuss in section 5. Another approach which includes explicitly the property of tense and at least half of the doctrine of presentism (i.e. the thesis that the future does not exist) is the 'causal set theory' of Rafael Sorkin, David Rideout and other prominent theorists, such as Faye Dowker.

   In our view, these approaches have a leg up on timeless multiverse quantum gravity approaches precisely because they are consistent with the veridicality of our experience of time and tense. In order to justify a claim so extraordinary as the unreality of time, evolution and illusoriness of our experience thereof, the timeless multiverse needs something beyond demonstration as a 'consistent' quantum gravity formalism.

4. *The atemporalist position agrees with ours with regard to the fact that different universes within a multiverse are now incommensurate with respect to a temporal measure*. Lastly, the atemporalist position may actually prove to be supportive of the fact of the finitude of the past, if its attempt to eliminate time as a measure should turn out to succeed only in part. Suppose, for example, that the concept of 'emergent time' turns out to be meaningful along with the objectivity of tense (we discuss this in section 4.6). The impact on supposed 'pre-Big Bang' cosmogonies can be enormous, if one can no longer stretch a meaningful time back 'through' the Big Bang 'singularity'. If universe phase *A* is incommensurate with respect to time with universe phase *B*, then it is just false to suggest that phase *A* precedes phase *B*. If we suppose that time is a meaningful local concept within *A* and *B*, then the implication is that both *A* and *B* have temporal origins. This is supportive of the finitude of the past and, hence, of the beginning of the universe.

4.2. Expanding universe

Pitts seems to think that the person who believes that current physical cosmogony lends significant support to (2) must be what he calls a 'GTR exceptionalist', that is, someone who tolerates singularities (and, perhaps, takes GTR to be the fundamental theory of gravity). A GTR exceptionalist would be a person who holds that our universe obeys the Hawking–Penrose singularity theorem (Hawking and Penrose, 1970). Hawking and Penrose assumed that gravity is always attractive (that is to say, their theorem dealt only with 'ordinary' types of matter). Just about 30 years ago, however, the idea of 'inflation' was introduced into cosmology as a resolution to several of the anomalies in the standard hot Big Bang picture, and a key characteristic of inflation is that it is driven by a type of energy that violates the energy condition of the Hawking–Penrose theorem. Thus it became a live question as to whether inflationary cosmogonies might accommodate a beginningless past. Perhaps we are living in a vacuum bubble that resides in a much larger (perhaps past and future fractal) structure of nested bubbles.

On this view, a small volume of 'false vacuum' may decay into the 'true vacuum' in which we live. This process, first suggested by Sidney Coleman and Frank De Luccia (Coleman and De Luccia, 1980) would identify our Big Bang with one of these quantum tunnelling 'decay' events. Thus the Big Bang would not represent an absolute beginning of space and time. Rather there would be a pre-Big Bang past.

What Pitts does not mention, however, is that singularity theorems have also evolved in the intervening 30 years. There has been a lively debate centred (mostly) around the role of energy conditions in these theorems. In 2003 Arvind Borde, Alan Guth and Alexander Vilenkin published a singularity theorem that was completely independent of energy conditions. Vilenkin explains:

> A remarkable thing about this theorem is its sweeping generality. We made no assumptions about the material content of the universe. We did not even assume that gravity is described by Einstein's equations. So, if Einstein's gravity requires some modification, our conclusion will still hold. The only assumption that we made was that the expansion rate of the universe never gets below some nonzero value, no matter how small. This assumption should certainly be satisfied in the inflating false vacuum. The conclusion is that past–eternal inflation without a beginning is impossible. (2006, p. 175)

In their formal paper, Borde, Guth and Vilenkin state that 'Our argument shows that null and time-like geodesics are, in general, past–incomplete in inflationary models, whether or not energy conditions hold, provided only that the averaged expansion condition $H_{av} > 0$ holds along these past-directed geodesics' (2003, p. 3).

Borde et al. suggest that timelike and null geodesics are redshifted in energy as they are stretched out in an expanding space. Looking backwards in time, they are blueshifted. The blueshift becomes infinite in a finite amount of proper time (or affine parameter for null geodesics). In GTR, the infinite blueshift suggests a singular condition.

The BGV conclusion applies beyond the inflationary model class. Any universe (including universes modelled by higher dimensional cosmology, pre-Big Bang cosmology and so forth) which, on average, expands has to connect, in a finite time, to a past boundary.[11] According to Vilenkin, this result is independent of modifications to Einstein's equations (which could be the result of 'low energy' corrections coming from a quantum gravity approach) or the particular characteristics of the proposed cosmogony. If the universe (or multiverse) expands (on average), then it has a beginning, period.

It is possible to object that the BGV theorem is a classical result lacking the rigor of a full quantum gravity approach. The objection is that BGV is based on metric tensors, which one could claim would not exist in a 'full' theory of quantum gravity.

As discussed in section 4.1, we believe the objection is without merit. We add that Borde, Guth and Vilenkin are not silent on the topic of quantum gravity. They indicate that their results demonstrate a boundary where a quantum gravity theory will be necessary to describe further the physics. Thus the classical dynamics governed by their theorem are adequate to describe the universe at 'low' energy states, including its past history roughly back to the Planck time. There is then a quantum gravity reckoning:

> What can lie beyond this boundary? Several possibilities have been discussed, one being that the boundary of the inflating region corresponds to the beginning of the Universe in a quantum nucleation event. The boundary is then a closed spacelike hypersurface which can be determined from the appropriate instanton.
>
> Whatever the possibilities for the boundary, it is clear that unless the averaged expansion condition can somehow be avoided for all past-directed geodesics, inflation alone is not sufficient to provide a complete description of the Universe, and some new physics is necessary in order to determine the correct conditions at the boundary. This is the chief result of our paper. (Borde, Guth and Vilenkin, 2003, p. 4)

As briefly mentioned earlier, a Feynman path-integral-based cosmogony such as the Hartle–Hawking 'no-boundary' (Hawking and Hartle 1983), or the Vilenkin 'tunneling from nothing' approach (Vilenkin 1982) can be interpreted as a creation-ex-nihilo[12] that would explain the boundary suggested by the BGV theorem as well as lend support to *kalam*'s second premise.

We discuss this quantum reckoning in section 5 of our paper. So either the boundary represents a quantum nucleation event or else (within the context of Pitts's original objection) we are left with the options discussed in sections 4.3–4.6 of this paper.

A possible further objection to the BGV theorem might draw upon Pitts's suggestion (2008, p. 681) that the Bach–Weyl approach to GTR does not assign lengths to curves. Demonstrating a past inextendible geodesic (i.e. one of finite length) is precisely how BGV purports to show that their boundary exists. Does the Bach–Weyl formalism give us, then, reason to doubt that objective scales exist?

We question whether Bach–Weyl should be taken as a realist approach to modelling. It may well be better understood instrumentally.[13] We note that some theorists seem to prefer a more modest use of conformal invariance. Roger Penrose's Cyclic Conformal Cosmogony

(see section 4.5), for example, uses exactly the Bach–Weyl formalism to model the boundary between consecutive oscillations. But Bach–Weyl requires all contents of the universe to be conformally invariant. Massless particles such as the photon and the graviton meet this criterion. In Bach–Weyl one splits a metric into a scale-free part and another part that measures space and time intervals. But massless particles (only) do not experience a passage of time. So if the universe contained only massless particles, one could argue (as Pitts does) that one could drop the part of the metric that represents intervals. But where *mass exists*, conformal invariance does not. So for all phases of the universe where mass exists (which in the real universe is pretty much every era after the Big Bang), one needs the full metric. Penrose thus splits his model into two epochs: one consistent with conformal invariance and one without it.

Similarly Gerard 't Hooft, an advocate of the Bach–Weyl approach, describes other necessary parts of a cosmogony that are not scale invariant and are thus problematic for a 'pure' Bach–Weyl methodology. In a paper defending the existence of space-time (as an objective feature of nature) at the expense of 'emergent time' theories he states,

> Describing matter in a $\hat{g}_{\mu\upsilon}$ metric will still be possible as long as we restrict ourselves to conformally invariant field theories, which may perhaps be not such a bad constraint when describing physics at the Planck scale. Of course that leaves us the question where Nature's mass terms come from, *but an even more urgent problem is to find the equations for the gravitational field itself, considering the fact that Newton's constant G_N is not scale-invariant at all. The Einstein–Hilbert action is not scale-invariant.* Here, we cannot use the Riemann curvature or its Ricci components, but the Weyl component is independent of ω, so that may somehow have to be used. ('t Hooft, 2009, p. 8; emphasis added)

Ultimately 't Hooft believes that scales are introduced in nature via the fundamental processes of information transfer. He sums up his paper by arguing, 'The density of this information flow may well define the Planck length locally, and with that all scales in Nature' (2009, p. 8). Thus he supports the idea that scales (time and space intervals) are objectively real features of nature.

It therefore seems to us that the mere fact of the Bach–Weyl approach (which is all that Pitts offers) proves little. It may describe possible worlds that are fully conformally invariant, but we are interested in a model that can represent the matter-filled universe we see around us.

What, then, has been the reception of the BGV theorem by the physics community? It, as well as its implications, is largely uncontroversial.[14] Instead, a new round of model building has ensued based on possible exceptions to this theorem. Four alternatives present themselves:

1. Infinitely contracting universe 'bouncing' into an expansion phase (average expansion < 0; example: de Sitter cosmogony).
2. Asymptotically static universe (average expansion $= 0$; example: Emergent model class).
3. Eternally cyclic universe (average expansion $= 0$; example: Penrose conformal cyclic cosmogony).
4. Reversal of the arrow of time (example: Aguirre–Gratton model).

Leaving (4) aside for the moment, let us consider universes that do not, on average, expand over their past histories.

4.3. Infinitely contracting universe

Is a pre-Big Bang infinite contraction (followed by a rebound at a near singular condition into our present day expansion) as envisioned in (1) a promising alternative? We suggest not. George Ellis comments on the problems that bedevil such an approach:

> The problems are related: first, initial conditions have to be set in an extremely special way at the start of the collapse phase in order that it is a Robertson–Walker universe collapsing; and these conditions have to be set in an acausal way (in the infinite past). It is possible, but a great deal of inexplicable fine tuning is taking place: how does the matter in widely separated causally disconnected places at the start of the universe know how to correlate its motions (and densities) so that they will come together correctly in a spatially homogeneous way in the future?
>
> Secondly, if one gets that right, the collapse phase is unstable, with perturbations increasing rapidly, so only a very fine-tuned collapse phase remains close to Robertson–Walker even if it started off so, and will be able to turn around as a whole (in general many black holes will form locally and collapse to a singularity).
>
> So, yes, it is possible, but who focused the collapse so well that it turns around nicely? (pers. comm., 25 January 2006)

First, then, such models encounter the significant problem of acausal fine-tuning. One asserts not just brute contingency but also a very

curious form of it. In the face of apparent fine-tuning, physicists usually prefer to offer some type of explanation. For example, cosmologists are avidly seeking an explanation for apparent fine-tuning of the parameters of the standard model (such as force-coupling constants) in the form of a multiverse or a superdeterministic Theory of Everything. Or one thinks of Guth's inflationary resolution of the horizon problem (past thermodynamic equilibrium). If we are going to give up explanation, then what was wrong with leaving cosmology as it was prior to 1980, namely, the standard hot Big Bang model (with the associated breakdown of physics at the singularity)?[15]

The second problem is that the collapse becomes chaotic as it approaches the singularity. This will produce a pre-expansion start condition that is known to be dramatically different from our actual Big Bang. This phenomenon is referred to as 'BKL chaos' after its discoverers (Belinsky, Khalatnikov and Lifshitz, 1970). The same problem will appear for all attempts at a past–eternal timeline that seek to introduce a pre-Big Bang phase that 'bounces' into the present expansion. In fact, the real implication of BKL may well be that it is physically impossible to 'bounce' *through* a 'singularity'.[16] So option (1) is unpromising.

4.4. Asymptotically static universe

An asymptotically static space, as envisioned in option (2), is one in which the average expansion rate of the universe over its history is equal to zero, since the expansion rate of the universe 'at' past infinity is zero (thus 'infinity' dominates any finite expansion phase, no matter how long). Hence, the universe, perhaps in the asymptotic past, is in a static state (neither expanding nor contracting). This feature of the model allows it to escape the BGV singularity theorem. Consider, for example, the GTR-based 'Emergent' model class of George Ellis et al. (Ellis, Murugan and Tsagas, 2004; Ellis and Maartens, 2004). It features two stages: an Einstein Static State (ESS) and an inflationary phase that leads to our present, dynamic, expanding universe.

The Einstein static universe itself was originally viewed as past eternal. But there are obvious problems with this model. The reason Einstein himself dropped it was its feature of unstable equilibrium. Although, in pure non-quantum GTR, one can consider a static state with worldlines that trace to negative infinite time, in reality we know that gravity is a quantum force. As Vilenkin notes, 'Small fluctuations in the size of the universe are inevitable according to the quantum theory, and thus Einstein's universe cannot remain in balance for an infinite time' (Vilenkin, 2006, p. 209). On the other hand, the current observable

universe is demonstrably *not* in a static state. So there would need to be at least two stages in such a model, a primordial ESS followed by an expansion phase. A quantum fluctuation would force a transition from ESS to an expanding universe. But this very mechanism implies that the *initial state* is not past eternal, since such a fluctuation will inevitably occur within a finite time.[17]

To preserve past eternality, one could claim that the static state is only an 'ideal point' asymptotically approached at past infinity. Some philosophers have expressed reservations, however, over the contrived nature of the past infinity featured in such an interpretation. The cosmological (and thermodynamic) arrows of time are so weak that they are indistinguishable from a 'timeless state'. For example, Rüdiger Vaas, characterizing the Emergent models as 'soft-bang/pseudo-beginning' in nature, views the asymptotic approach toward ESS as a mathematical artefact (Vaas, 2004, p. 18).

A further problem with these models is the implied fine-tuning of the initial state. One way to look at this is to imagine starting in the present and extrapolating the past history of the universe. Will the resultant evolution, with high probability, produce an ESS state? No, such an outcome is improbable. Ellis is sensitive to the fine-tuning problem but thinks his approach worth pursuing in the absence of a quantum gravitational resolution (Ellis and Maartens, 2004, p. 228).

Subsequent to his initial work on Emergent models, which were based solely on GTR, Ellis along with his colleagues integrated a quantum gravity approach into their models. Ellis's colleague David Mulryne elaborates: 'The importance of the [equilibrium state as described by quantum gravity] is that, in contrast to the [ESS] solution present in GTR, slight perturbations do not result in an exponential divergence from the static universe, but lead instead to oscillations about it' (Mulryne et al., 2005, p. 6). The ESS is now viewed as a 'low-energy' solution of loop quantum gravity (LQG) to make the Einstein state stable against perturbations of a limited size. LQG theorist Martin Bojowald explains how Mulryne et al. use the mechanism of perturbation to cause an initially stable oscillation to escape to full-blown inflation:

> Static solutions do not evolve, and so are clearly ill-suited as a model for the Universe. *But by introducing a perturbation to a static solution, one can slightly change it and thereby start a more interesting history*. Unfortunately, the classical solution (ESS) is unstable: any disturbance grows rapidly, leaving little of the initial state behind. The insight of Mulryne and colleagues is that quantum effects could

> supply all the necessary ingredients where classical solutions do not. Within the framework of loop quantum gravity, repulsion also implies static solutions at small size, but these – in contrast to the classical case – are stable. *According to the authors' model, perturbing such a state leads to small cycles of interchanging expansion and contraction.* During this process, matter will evolve slowly, and the cycles will gradually change their behavior. By itself, this perpetual recurrence and incremental change seems to lack the spark necessary for so momentous an event as the birth of the Universe. And indeed, Mulryne and colleagues identify one final theoretical ingredient that lights this spark: mediated through repulsive effects, *potential energy is gradually pushed into the matter during its slow evolution. At the point when potential energy starts to dominate kinetic energy, the mundane cycling is broken by a sudden, dramatic inflationary explosion – the emergent Universe*. (Bojowald, 2005, pp. 920–1; our emphasis)

Should the past be judged infinite in such an approach? There *is* a claim by its authors to possible past eternality: 'The universe undergoes a series of non-singular oscillations in a (possibly) past–eternal phase with the field evolving monotonically along the potential' (Mulryne et al., 2005, p. 6).

Similar to the earlier Emergent models of Ellis et al., which proposed an ideal ESS point 'reached' in the infinite past, the model of Mulryne et al. proposes that any point which is close to the stable equilibrium point could serve as an 'initial' state for the model. The size of the universe, then, oscillates about this initial state, always maintaining a non-zero positive value. This oscillation has an amplitude which is associated with a maximum size for the universe (per oscillation) which we will call $\mathbf{amax}_i$ (see Bojowald and Tavakol, 2008a, fig. 2, and Mulryne et al., 2005, fig. 7). In the absence of the self-interaction property of the scalar field φ, this condition is said to be perpetual. But the potential energy in the system must grow, forcing **amax** to grow with each subsequent cycle. This is said to introduce a cosmological arrow of time, which is based on the evolutionary increase in **amax**. We will take the equation **amax(t)** to be fundamental; the potential $V(\varphi)$ and scalar field $\varphi(t)$ are reverse engineered to produce it (see Ellis, Murugan and Tsagas, 2004, section II text and fig. 1, and Mulryne et al., 2005, fig. 5).[18]

In the limit $\mathbf{t} \to -\infty$, it is desired that the change in **amax** (call it $\mathbf{\Delta a}_{\mathbf{evolution}}$) between cycles asymptotes to zero. Note that the interval in time between a cycle with maximum amplitude $\mathbf{amax}_i + \mathbf{\Delta a}$ (with arbitrarily small Δa) and the present is finite; this is shown by the numerical

analysis in Mulryne et al., 2005. So if the initial process of reaching $\mathbf{amax}_i + \mathbf{\Delta a}$ has a finite timeline, the full timeline to the present is finite.

Now, consider a quantum fluctuation of the scale factor, $\mathbf{\Delta a}_{\mathbf{fluctuation}}$. This is a separate, instantaneous, mechanism whereby the universe may reach a larger size. For the universe at **amax** (Bojowald and Tavakol, 2008a, fig. 2 and equations 32 and 33), and in the past asymptotic limit of an unsqueezed quantum state,[19] the fluctuation is roughly proportional to **amax**. So looking to the infinite past, $\mathbf{\Delta a}_{\mathbf{fluctuation}}$ will be $>> \mathbf{\Delta a}_{\mathbf{evolution}}$.

This suggests two things. First, as one looks to the past, the (typical) size of fluctuations of the scale factor eventually exceeds and dominates the evolutionary increase (in **amax**) between cycles. As Vaas might describe it, there is no cosmological arrow of time in this region. Second, it does not seem reasonable to assert that the universe will need an infinite amount of time to grow $\mathbf{amax}_i$ to amplitude $\mathbf{amax}_i + \mathbf{\Delta a}$ (where $\mathbf{\Delta a}$ can be arbitrarily small). The probability should be unity that it will reach that goal by instantaneous, discrete transition in finite time, as opposed to continuously evolutionary growth over infinite time.

Furthermore, since the early phase of this model features an oscillating universe, it also encounters an entropy objection to past eternality (see section 4.5). It seems, then, that the past eternal nature of this model is a mathematical artefact. One need not reject the model; one merely notes that there must a temporal origin of some sort in the finite past.

In sum, here is a counterexample to Pitts's notion that a past lacking a singularity must of necessity be eternal.[20] The Emergent models that feature (only) a GTR based ESS are metastable.[21] An improvement to the Emergent model resolves this instability by providing quantum-based modifications to the Friedman equation. As the authors say, this provides 'partial' ameliorization of the metastability objection and the fine-tuning objection (Mulryne et al., 2005, abstract). But the nature of quantum fluctuations demands that one cannot have an infinite timeline during which an arbitrarily small, continuous evolution in scale factor occurs. It seems to us this objection would apply, generally, to asymptotically static constructions.[22]

So far, then, we suggest that there are good arguments to believe that contemporary pre-Big Bang universe models that contract, are static or expand (on average) either imply a beginning or are untenable. The expanding models obey a newly discovered singularity theorem (BGV) and likely have a beginning of a quantum nature. The contracting models featured in alternative 1 seem to be unpromising and add nothing

to the standard hot Big Bang model from the standpoint of *explanation*. The static models featured in alternative 2 imply a beginning but do not necessarily describe how that state came to be.

4.5. Eternally cyclic universe

What, then, about cyclic models, as envisioned in option 3 above? According to these models, the universe goes through a cycle in which it grows from zero (or near-zero) size to a maximum and then contracts back to its starting condition. The universe itself is periodic, in the sense that it undergoes many such cycles, perhaps an infinite number. The average expansion of the universe would be zero in a 'pure' cyclic model, since cycle by cycle, the universe always undergoes precisely equal amounts of expansion and contraction. Hence, a cyclic model evades the BGV Theorem.

Cyclic models face a well-known thermodynamic problem. As Vilenkin notes, 'A truly cyclic universe has a problem with entropy increase: it should have reached thermodynamic equilibrium by now' (pers. comm., 19 January 2008). Our observation of the present day universe indicates that we are not at a condition of thermodynamic equilibrium – a good thing for us, as life requires nonequilibrium conditions in order to exist. As one looks into the past, the size of each cycle is also thought to decrease (due to radiation effect on entropy; see Barrow and Dabrowski, 1995). Eventually a 'Planck entropy' could be reached (the minimum currency of entropy exchange), which would preclude the existence of still earlier cycles. A problem similar to that encountered for an asymptotically static cosmogony can also occur. *If one is attempting a realist ontology*, the backward evolution of a cyclic model must stop once quantum fluctuations in volume are of the same size as the volume expectation value. One can continue pursuit of the problem (the classical background disappears) in a full quantum approach (pers. comm., Martin Bojowald, 15 April 2010). If this means that time 'disappears' or that the arrow of time reverses, the implications are discussed in section 4.6.

Cosmologists Thomas Banks and Willy Fischler contend that a contracting space filled with quantum fields will have an 'ergodic' property as the space shrinks. Its fields become highly excited as one approaches the end of contraction, and these fields will produce chaotic fluctuations. Spontaneously created matter with a different equation of state will dominate the energy density. That, and the inhomogeneity of the fluctuations, will prevent cycling. Banks and Fischler even suggest that the fields will spontaneously produce a dense 'fluid' of black holes

leading to a condition they call a 'Black Crunch' (Banks and Fischler, 2002) for arbitrary states approaching full contraction.[23]

Martin Bojowald and collaborators (Bojowald, Maartens and Singh, 2004; Bojowald, 2006; Bojowald and Tavakol, 2008a, 2008b) have done a significant amount of work on building contemporary cyclic models to address these difficulties. The LQG approach does seem to resolve the singularity admirably in its initial tests against simple (homogeneous) models (Ashtekar, Pawlowski and Singh, 2006; Bojowald, 2007), although it is unclear if this approach solves all the problems brought up by opponents such as Banks and Fischler.[24]

In 2006, Bojowald suggested the possibility that the recollapse phase of each oscillation was entropy-reducing (see Craig and Sinclair, 2009, p. 172). Perhaps the universe does not cumulatively gain entropy cycle to cycle. But if so, this could lead to a different problem in that it is possible to interpret an *entropy reducing* recollapse in terms of a reversed arrow of time; thus the 'recollapse' is really another expanding universe. We shall take up this question again in section 4.6.

In 2008, Martin Bojowald and Reza Tavakol entertained the opposite suggestion; entropy is monotonically gained from cycle to cycle. Quantum squeezing of the conjugate variable pair V (volume) and P (momentum) of the universe was identified as an entropy contributor without classical parallel. Interestingly, the quantum gravity model now displays the same behaviour with respect to entropy gain that Barrow and Dabrowski (1995) note for classical models in their paper:

> An interesting question is whether in a cyclic model one generically expects to have a finite or an infinite number of past cycles. The problem with the finite case is that it does not resolve the origin question. In the emergent scenarios, as well as some other such models, the universe is assumed to have undergone an infinite number of past cycles so as to remove the question of the origin. In that case any given cycle would have an infinite number of precursors and generically we therefore have to expect the current state to be squeezed . . . . The question then is how the squeezing in a generic cycle is determined. *If each cycle produces the same amount of squeezing, a generic cycle would have infinitely squeezed states, which could not be semiclassical . . . For growing cycles, as in the emergent scenario, the change in squeezing is initially small and approaches zero for cycles in the infinitely distant past.* Depending on the precise scenario, the sum of all squeezing contributions may converge, such that a finite value

> results for a generic cycle. (Bojowald and Tavakol, 2008a, p. 8; our emphasis)

Similarly Aguirre comments:

> For an observer along that worldline to perceive an arrow of time (AOT) it must see local net entropy generation. To the past, it sees entropy destruction. In a finite neighborhood, there is finite entropy, so going far enough 'to the past' the entropy must (a) asymptote to a constant, (b) start to increase again, or (c) become ill-defined as a singularity is encountered. (Aguirre, 2007, p. 30)

So avoiding a reversal of the AOT requires entropy to asymptote to a constant in the infinite past. In a quantum model, the squeezing of states is going to be a factor in accomplishing this. In their paper Bojowald and Tavakol suggest both that a special 'unsqueezed' state represents the asymptote but also that it is problematic to assert that such a state existed. Bojowald (pers. comm., 15 April 2010) has indicated to us that there were further problems since discovered with his and Tavakol's assumptions on monotonicity of variables. So, as of this writing, the problem of structuring entropy to be infinitely subdividable remains.

Returning to the general discussion, suppose that total entropy can be kept constant with cumulative cycles and that contracting phases are not reinterpreted as having a reversed arrow of time. There remains the issue of dark energy, which may have the potential to halt the cycling and induce an open-ended expansion. The current empirically observed dark energy effect, for example, appears adequate to produce an open-ended, accelerated expansion. This result would be definitive if the dark energy is of the form of a cosmological constant (that is to say, if its value is independent of space and time (see Barrow and Dabrowski, 1995). Indeed, open-ended, accelerating expansion does appear to be the fate of the present day universe (Overbye, 2006). If an entropy gain (cycle-to-cycle) is denied, then the amplitude of the oscillation (maximum scale factor experienced by the universe) is constant. But then one could never have more than one 'cycle', for the cosmological constant would lead to open-ended expansion *the first time*. Hence, the initial cosmological singularity (our Big Bang) *would* represent an absolute beginning.

There are some genuinely exotic cyclic models that purport to escape these problems (see Craig and Sinclair, 2009 for an analysis of various cyclic models current in contemporary cosmological discussions).

A fascinating attempt to counter the entropy and dark energy problems mentioned above is Roger Penrose's Cyclic Conformal Cosmogony (Penrose, 2006). Penrose suggests a solution to the entropy problem by saying that the initial 'singularity' is the *same thing* as the open-ended de Sitter-like expansion that our universe seems about to experience. Their mathematical equivalence is demonstrated through an appropriate conformal transformation of one state into another. Penrose explains how we are to think about this situation:

> Physically, we may think that again in the very remote future, the universe 'forgets' time in the sense that there is no way to build a clock with just conformally invariant material. This is related to the fact that massless particles, in relativity theory, do not experience any passage of time.... With conformal invariance both in the remote future and at the Big-Bang origin, we can try to argue that the two situations are *physically identical*, so the remote future of one phase of the universe becomes the Big Bang of the next. (Penrose, 2006, p. 2761; emphasis original)

Penrose admits that his view is heterodox in that, for this envisioned scenario to obtain, all massive fermions and massive, charged particles must disappear to radiation, including, for example, free electrons. He concedes that there is (currently) no justification for positing this.

The Cyclic Conformal Cosmogony is based on the Weyl Curvature Hypothesis (WCH) and Paul Tod's implementation of this idea within GTR (Tod, 2003). WCH is defined as follows:

> Weyl curvature is the kind of curvature whose effect on matter is of a distorting or tidal nature, rather than the volume-reducing one of material sources.... The physical conjecture that I refer to as the Weyl curvature hypothesis asserts that (in some appropriate sense) the Weyl curvature is constrained to be zero (or at least very small) at initial singularities, in the actual physical universe. (Penrose, 2005, pp. 765, 768)

Penrose describes the mathematical technique necessary to stitch a singularity to a maximally extended de Sitter expansion:

> Tod's formulation of WCH is the hypothesis that we can adjoin a (past-spacelike) hypersurface boundary to space–time in which the conformal geometry can be mathematically extended smoothly

> through it, to the past side of this boundary. This amounts to 'stretching' the metric by a conformal factor Ω which becomes *infinite* at the Big Bang singularity, so that we get a smooth metric $\hat{g}_{ab}$ which actually extends across this boundary. So far, we regard the conformal 'space–time' prior to the Big Bang as a mathematical fiction, introduced solely in order to formulate WCH in a mathematically neat way. However, my 'outrageous' proposal (4) is to take this mathematical fiction seriously as something *physically real*. (Penrose, 2006, p. 2761; emphasis original)

A potential failing of this approach is the supposed correspondence between Weyl curvature and entropy. The correspondence seems clear enough when one is considering the structure of the initial Big Bang singularity, given its vanishingly small entropy state. But while the de Sitter-like end state of the universe also minimizes Weyl curvature, its entropy is *maximized*. Like a black hole obeying the Hawking–Bekenstein entropy law, a de Sitter space has a cosmological horizon with entropy proportional to its area. It is generally believed that this state represents the maximum entropy that can fit within the horizon.

Penrose has recently chosen to regard the entropy of the cosmological horizon as spurious and to invoke non-unitary loss of information in black holes in order to equalize the (vanishingly small) entropy at the boundary (Penrose, 2009, p. 15). Penrose attributes the large entropy at late universe times (but before significant decay of the universe's black hole population) to degrees of freedom internal to the black holes. He then suggests that in CCC the universe's entropy is 'renormalized' so that we can discount the entropy contribution from the horizon when all black holes have evaporated.

On the other hand, it seems, Penrose notwithstanding, that the entropy of the cosmological horizon *must* have physical meaning, since the entropy of the de Sitter-like system outside a black hole must be higher than that of the black hole itself. For if it were not, then the physics of black hole decay (upon which Penrose's scenario depends) would not work properly. Black hole decay is actually a dynamic system which is the sum of the energy lost by Hawking radiation plus the energy gained by absorption of local matter created by thermal fluctuations due to the de Sitter Gibbons–Hawking temperature. The mere fact of these thermal fluctuations suggests that the entropy of the cosmological horizon is a real physical manifestation as opposed to an accounting gimmick.

As a thought experiment, one can imagine what happens as more and more external matter (over and above that added by external environment thermal fluctuations) is added to a black hole. Is there a maximum size? As the black hole gets larger, its temperature drops, approaching that of the cosmic background. In fact, at a certain point the black hole itself becomes physically indistinguishable from a cosmic horizon and can be identified as such. At no point does it become unreal. It therefore seems unwarranted to embrace Penrose's latest (2009) position, and very few physicists have been persuaded by Penrose's non-unitary brand of quantum physics.

In sum, while Weyl curvature is the same between the two states that Penrose wishes to say are identical, the entropy is not. It seems to us, then, that the two states cannot be identical. Hence, we question CCC's viability.

It seems, then, that cyclic models do not avert the conclusion that time had a beginning. Thermodynamic objections to cyclic models, 'classical' anticipations of which were seen in Richard Tolman's early work in the 1930s (Tolman, 1934), still seem to find purchase. Entropy should increase from cycle to cycle, implying that (1) we should now be in thermodynamic 'heat death' (which we are not), and (2) the maximum size (scale factor) of the universe should grow with each cycle. This large scale 'classical' behaviour likely ends when (looking into the past) the cycle size shrinks to Planckian physics. If entropy gain is denied, then the empirical observation of a cosmological constant of sufficient size to lead to an open-ended expansion rather than a future contraction (this does seem to be the fate of our universe) *also* implies that there can have been no previous cycle. If our future does include a return to a singular condition, then, as Banks asserts (recalling our earlier discussion of BKL chaos), 'It seems silly to imagine that, even if (a future singular condition for the universe) is followed by a re-expansion, that one would start that expansion with a low entropy initial state, or that one had any control over the initial state at all' (pers. comm., 17 October 2008).

4.6. Reversing time's arrow

There is a longstanding debate in cosmological circles as to how to interpret universe phases that would be entropy-reducing. There is no necessity (in GTR) for the cosmological and thermodynamic arrows of time to align with each another, although they seem to do so in the present era of our universe. This debate has consequences for discussion

of a cosmic origin. For example, in inflationary models, the BGV theorem seems to bar a beginningless past. But suppose one takes a model that features a contraction, a bounce through a near-singular state, and then an expansion. Such a model faces fine-tuning objections such as that discussed in section 4.1. Yet it is possible to (re)consider the contracting phase as having a reversed arrow of time. The boundary that formerly represented the 'bounce' will now bisect two symmetric, expanding universes on either side.

On such a view space-time is extendible beyond the boundary for an observer, looking backwards in time, from either side. Does this imply past eternality? Cosmologists Anthony Aguirre and Steve Gratton think so. They have employed the concept to create an 'eternal steady state inflation model' (Aguirre and Gratton, 2002). To visualize their idea, start with the standard foliation of de Sitter space as looking something like an hourglass. The width of the glass represents the scale factor of the universe. Vertical height (pointed upward) is the time. Thus the space shrinks to a minimum radius and then rebounds into an expansion. On this understanding, one cannot possibly have 'past eternal inflation'. Borde, Guth and Vilenkin explain:

> The intuitive reason why de Sitter inflation cannot be past–eternal is that, in the full de Sitter space, exponential expansion is preceded by exponential contraction. Such a contracting phase is not part of standard inflationary models, and does not appear to be consistent with the physics of inflation. If thermalized regions were able to form all the way to past infinity in the contracting space-time, the whole universe would have been thermalized before inflationary expansion could begin. (Borde, Guth and Vilenkin, 2003, p. 1)

To obtain a different outcome, Aguirre and Gratton choose to interpret de Sitter space according to a procedure first suggested by Erwin Schrödinger in 1956. Utilizing 'elliptic de Sitter space', they remap the hourglass by splitting it in two along a line from the top left to the bottom right. For all intents and purposes, these are two separate universes. Maulik Parikh explains how this looks to antipodal observers on opposite sides of the divide:

> The antipodal map, $X^I \rightarrow -X^I$, (where 'I' is a dimensional index) changes the sign of the time coordinate of the embedding space, and also that of the direction of time in de Sitter space. The resulting quotient space, dS/Z2, *is as a consequence not time-orientable*: although

> one can locally distinguish past and future, *there is no global direction of time*. This fact clearly changes many standard notions about space and time that we are accustomed to. For instance, it is impossible to choose a Cauchy surface for elliptic de Sitter space that divides space-time into a future and a past region. (Parikh et al., 2003, p. 6; our emphasis)

Aguirre and Gratton explain the relationship between the two disconnected regions:

> In essence, this construction partitions the full de Sitter space-time into a self-consistent set of two noncommunicating (steady state) universes. An observer in region I does not see anything in its past light cone from an observer in region II because that other observer cannot signal into its past, and vice-versa. *Seen in this way the boundary condition forbidding physical particles from following geodesics across (the boundary) – into one universe is in no way strange or unreasonable, as it follows directly from the forbidding of causality violations in the other universe.* (One could similarly partition de Sitter space-time by any non-timelike boundary B away from which time flows.) . . .
>
> Without the identification, the space-time manifold is time-orientable in the mathematical sense that it is possible to continuously divide non-spacelike vectors into two classes which can be labeled 'future' and 'past.' In our construction these labels will only correspond to the physical (arrow of time) in one of the two regions. With the identification the space-time is still a manifold but is not mathematically time-orientable. The physical (arrow of time) is, however, still well-defined and no physical observer will see it reverse. (Aguirre and Gratton, 2002, p. 3; our emphasis)

As can be seen, neither region of the remapped space stands in a relation of *earlier than* to the other. There is no global definition of time. As Parikh notes above, it is impossible to determine a future and past region (when considering the whole hourglass). A matter of some importance is that there is no communication; no *causality* by definition from one region to the other. Thus one can say that there is no eternal past that evolved into our present. So Aguirre and Gratton's use of the term 'past eternal' seems to be idiosyncratic in that they construe one's ability to draw 'complete' geodesics (meaning 'from one end of the space-time to the other', even though no observer or causal chain could traverse them) from end to end on a graph to be a sufficient

condition of 'eternal'. What one really has in this model are two separate universes that trace their origins to a past boundary (the very boundary that Borde, Guth and Vilenkin have demonstrated).

Aguirre and Gratton conclude: 'We suspect that a construction like that proposed here may be necessary in any reasonable model for an eternal universe that avoids a beginning of time' (Aguirre and Gratton, 2002, p. 6). We suggest, on the contrary, that if they are right, then a beginning of time *has* been demonstrated. Rather than a past eternal *universe*, one has a past finite *multiverse*. This model is a perfectly viable example of a cosmological model that averts the singularity theorems yet still 'begins to exist', despite a technical past extendibility through the boundary.

We may resume here our discussion of cyclic models in section 4.5. Martin Bojowald has recently argued that the bounce events of a cyclic model actually represent a scenario similar to the Aguirre–Gratton 'double Big Bang' described above. Measurements of a state of the universe 'after' a bounce cannot realistically be used to derive knowledge of the state of the universe 'before' the bounce. Thus the model contains 'cosmic forgetfulness' as a salient feature. Similarly, an observer just 'before' a bounce would find that attempts at predicting the post-bounce future suffer from cosmic forgetfulness as well. Bojowald interprets the situation as follows:

> The kind of cosmic forgetfulness realized in this model provides an orientation of time, telling us not only which of the properties before the Big Bang can be forgotten, but also what direction 'before the Big Bang' is. An observer after the bounce would be unable to reconstruct the full state before the bounce, but could easily predict the future development toward larger volume. This arrow agrees with the standard notion.
>
> Now asking how an observer before the Big Bang would experience the same situation, the answer is also clear: such an observer would be unable to determine the precise state at larger values of φ (which is a matter field used in the role of a clock) beyond the bounce, but could easily extrapolate the state to smaller values of φ (the direction away from the 'bounce'). The state at smaller values of φ can be predicted, while the state at large values of φ is forgotten once the bounce is penetrated. *Since one cannot forget the future, such an observer must be attributed a reversed arrow of time*, pointing toward smaller φ. *At the bounce, two arrows would emerge pointing in opposite directions as far as φ is concerned*. In this sense, the model resembles [Aguirre–Gratton and

> similar suggestions from two other modelling teams using different methods] . . .
>
> Taking the simplest models of loop quantum cosmology at face value is often seen as suggesting the big bang transition to be viewed as a smooth bounce, as one further element not just in a long history of the universe itself but also in a long history of bouncing cosmological models. Some indications, however, suggest otherwise. The bloomy scenario of loop quantum cosmology may well be this: a universe whose time-reversed prehistory we cannot access but which we grasp in the form of initial conditions it provides for our accessible part; a pseudo-beginning; an orphan universe, shown the rear-end by whatever preceded (and possibly created) it. (Bojowald, 2009, p. 15; our emphasis)

Recall we had earlier discussed the breakdown of models featuring a classical background plus quantum fluctuations. A hypothetical resolution is to go to a full quantum gravity model. Would this produce past eternality? Bojowald suggests it produces a reversed arrow of time on the other side of the boundary. But, as the moments on one side of the boundary are in no sense earlier than the moments on the other side, there would be no infinite past.

H. D. Zeh, in the same volume in which Bojowald's essay appears, gives independent justification for accepting the interpretation of the boundary as an origin as opposed to a 'bounce'. The Wheeler–DeWitt (WDW) wave function (upon which loop quantum gravity is based) would be symmetric on both sides of a 'bounce'. This suggests that what one really has is a 'double Big Bang' coming from the same creation event. Zeh also suggests that the semi-classical approximation which defines time in a WDW approach cannot, with validity, be extended 'through' a bounce. Lastly, the low entropy initial state is common for all branches of the wave function.

> According to loop quantum cosmology, the Wheeler–DeWitt equation (in this theory replaced by a difference equation with respect to (scale factor) a) can be continued through $a = 0$ to negative values of a. The configuration space of three-geometries is in this way duplicated by letting the volume measure assume negative values (turning space 'inside out' while going through $a = 0$). Since the Hamiltonian does not depend on the newly invented sign of a, however, the Wheeler–DeWitt wave function must be expected to be symmetric under this parity transformation, too. Its continuation would then have to be interpreted as an added superposition of

> other physically *expanding* universes. Since the WKB times, which represent classical time, can *not* be continued through $a=0$, the interpretation of negative values of a as representing pre-big-bang times is highly questionable. The fundamental arrow, including its consequence of decoherence outside the validity of a WKB approximation, must depend on some low entropy 'initial' condition in a for all other ('spacelike') degrees of freedom that occur as physical arguments of the Wheeler–DeWitt wave function. It would be hard to understand how the low entropy state at $a=0$ could have been 'preceded' by an even lower entropy at $a<0$ in order to avoid a reversal of the thermodynamical arrow in the classical picture of an oscillating universe. (Zeh, 2009, pp. 11–12; author's emphasis)

On this view, a loop quantum approach predicts an origin out of which a multiverse arises. Continuation through the 'bounce' boundary actually represents other physically expanding universes. Zeh and Claus Kiefer arrive at a similar conclusion arguing directly from the Wheeler–DeWitt equation within the formalism of quantum geometrodynamics.

> One might wonder what happens in the case of models which classically describe bouncing cosmologies: the Universe would then undergo many, perhaps infinite, cycles of expansion and recollapse. What would happen with the entropy in these cases? If the entropy were indeed correlated with the scale factor, as the scenario discussed above suggests, the arrow of time would not continue through a turning point. The bouncing models would thus make no sense in quantum cosmology; one would only have branches of the wave function in which the arrow would point from small to large universe and where time would end when approaching a classical turning point. (Kiefer, 2009, p. 9)

> *The different quasiclassical branches of the wave function which are connected by 'quantum scattering' at the turning point should rather be interpreted as all representing different expanding universes*, which disappear at the turning point by means of destructive interference (similar to their coming into existence as separate Everett branches from a symmetric initial state at the big bang). (Kiefer and Zeh, 1994, p. 4152, our emphasis)

Here, Wheeler–DeWitt physics (with a boundary condition of a vanishing wavefunction) seems to predict the creation *ex nihilo* of an Everett multiverse as opposed to a recollapse.

Recalling Penrose's model, suppose that we grant his reversal in entropy as one looks to the universe's far future. One may just as well interpret his model as having a reversed arrow of time when looking to the 'past' of a Big Bang singularity, so as to obey a second law of thermodynamics. As such, it would display precisely the behaviour cited above by the Aguirre–Gratton, Bojowald and Kiefer–Zeh approaches. Physically, the question is whether, in our universe, various arrows of time (thermodynamic, cosmological, electromagnetic, psychological) align. Philosophically, the question is whether there is an underlying metaphysical time which the physical times manifest in various ways.[25] Extendibility, by this understanding, would be a technical artefact rather than an indication of past eternality.

Positing the beginning of classical time at a past boundary may have a dramatic implication first pointed out by philosopher Quentin Smith. Smith's comments were aimed at Hawking's interpretation of the imaginary time featured in the Hartle–Hawking quantum gravity cosmogony. Hawking was attempting to posit an 'imaginary' time phase for the universe that transitioned into a later, 'normal' time phase. Says Smith,

> such an interpretation is implicitly logically self-contradictory (in featuring a 'time' axis that starts off just like a space dimension, but then transforms to its familiar properties). The problem appears in the statement that the four-dimensional space joins on to the real (Lorentzian) space-time 'once' (*i.e.*, after) the quantum smearing effects subside: 'The question then arises as to the geometry of the four-dimensional space which has to somehow smoothly join onto the more familiar space-time once the smearing effects subside.' If the four-dimensional space does not possess a real time value, how can it stand in relation to the four dimensional space-time of being earlier than it? If the four dimensional space is not in real (Lorentzian) time, then it is not really earlier than, later than, or simultaneous with the four dimensional space-time manifold. Accordingly, it is false that the 4-sphere joins onto the familiar space-time once (*i.e.*, after in real time) the quantum effects dissipate. (1993, p. 318)

Smith later showed (1997, 2000) that these two (actually three; there is a transition region) 'stages' of the H–H model could coherently coexist in the same model only if they stood in a more primitive 'topological' relation,[26] rather than a temporal one. Thus, if two stages of a cosmogony did not stand on *one and the same* time axis, then it would be false to ascribe to them an *earlier than* or *later than* relation, even though

extendibility existed between the relevant universe phases. Recall, for example, that Penrose denies a real time value to his boundary as a physical necessity. When Penrose says, 'the remote future of one phase of the universe becomes the Big Bang of the next', his statement is arguably incoherent in the context of his model.

It seems to us Smith's insight applies in general to cosmological models whether they are classical, general relativistic models (which thus include singularities) or quantum gravity models (which include 'strong quantum regions'). The 'first' phase (which, alternatively, could be a 'transition' boundary between different classical or semi-classical space-times) by definition does not possess a time value. As Smith points out, because a model phase that lacks a real time value is not before, simultaneous with, or after any other universe state (with respect to a phase that does have real time values) it is not correct to say that the universe evolved from some state 'at' the boundary into some state with a real time value.

Very similar language is used by loop quantum theorists themselves.[27] On Smith's view (similarly, Ashtekar et al.), it seems correct to view this era as a primitive manifold with topological relations only between the 'cycles'. The cycles in the 'earliest' era seem to be a set of disconnected space-times which lack temporal ordering. One could just as well razor[28] these out. One is left with a space-time with a finite past timeline and no first moment, which grows into a macroscopic universe similar to ours.

Thus it seems 'singularities' are resolved, or not. If so, one has a pre-Big Bang model and faces issues such as a disappearing arrow of time, a reversed arrow of time, and/or entropy problems. If not (meaning time 'dissolves'), then there are no 'pre-Big Bang' phases, since those phases are incommensurate with respect to our FRW-like universe.

One recalls Pitts's insistence on a first moment as a necessary condition of a beginning of time (2008, p. 677). Precisely the opposite situation may obtain, generically, in models of the types discussed, despite the technical extendibility of geodesics through the boundary. We suggest that this situation represents a creation *ex nihilo* just as surely as the Big Bang singularity in standard cosmology itself represents an absolute beginning.[29]

In summary, then, it seems to us that a strong case for a beginning of the universe can be made with respect to pre-Big Bang cosmogonies (only some of which involve non-singular models), so that (9) is dispensable so far as premise (2) of the *kalam* cosmological argument is concerned, regardless of how critical it might be to Pitts's imagined singularity argument.

5. The Beginning of the Universe

But what, we may ask, is the nature of the universe's beginning if it is not a GTR-type singularity? The most widely discussed framework for addressing this problem is the Vilenkin 'tunneling from nothing' or the Hartle–Hawking 'no boundary' model.[30] Vilenkin observes,

> Many people suspected that in order to understand what actually happened in the beginning, we should treat the universe quantum-mechanically and describe it by a wave function rather than by a classical space-time. This quantum approach to cosmology was initiated by DeWitt and Misner, and after a somewhat slow start received wide recognition in the last two decades or so. *The picture that has emerged from this line of development is that a small closed universe can spontaneously nucleate out of nothing, where by 'nothing' I mean a state with no classical space and time.* The cosmological wave function can be used to calculate the probability distribution for the initial configurations of the nucleating universes. Once the universe nucleates, it is expected to go through a period of inflation, driven by the energy of a false vacuum. The vacuum energy is eventually thermalized, inflation ends, and from then on the universe follows the standard hot cosmological scenario. (2002, p. 2; our emphasis)

Hartle and Hawking also claim that their cosmogonic model can be interpreted in such a way that on that interpretation the universe came into being out of 'nothing':

> One can interpret the functional integral over all compact four-geometries bounded by a given three-geometry as giving the amplitude for that three-geometry to arise from a zero three geometry; that is, a single point. In other words, the ground state is the probability for the Universe to appear from nothing. (1983, p. 2961)

On these models the universe clearly has a beginning; but what does it mean to 'nucleate out of nothing?' It seems to us that there are four possible ways of interpreting this notion:

1. The 'initial' null topology[31] represents literal nothingness. One could imagine that 'nothing' as employed by Hartle–Hawking has exactly the relevant meaning for the *kalam* cosmological argument, namely, the assertion of nonbeing. If this is what is meant here, then models of this sort feature an unambiguous creation *ex nihilo*. The second

premise of the *kalam* cosmological argument is upheld, and the debate shifts to the first premise.

2. The 'initial' null topology is to be construed purely instrumentally. Hawking himself seems to give good grounds for treating Hartle and his proposal as a FAPP (i.e. 'for all practical purposes') approach only. The H–H formalism considers a superposition of all possible quantum geometries, with our universe emerging as the most probable, and uses Richard Feynman's 'sum over histories' approach to quantum mechanics to predict universe observables. This formalism treats time, in the quantum gravity era, as a fourth spatial dimension. But in his collaboration with Roger Penrose, *The Nature of Space and Time*, Hawking employs the same mathematical approach (analytic continuation) to describe pair production of electron/positron pairs in a strong electric field (1996, p. 54). This is a standard mathematical technique sometimes used when complex analytic functions are better behaved in a certain domain than their real counterparts. It does not imply ontological commitment to the alternative description, however. It seems to us that given the unintelligibility of the 'imaginary time' region in these models, it is most reasonable to treat this regime as a useful fiction. Thus, the model serves only to reinforce the second premise of the *kalam* cosmological argument.
3. The 'initial' null topology is literally something. Attempts to take the imaginary time interpretation realistically can run into problems similar to those Smith pointed out (1993, 1997). In fact, more recent work using computer simulations of the evolution of space-time suggests that classical space-time cannot have emerged from a Euclidean 4-space such as is envisioned in the H–H model. Jerzy Jurkiewicz, Renate Loll and Jan Ambjorn explain:

> In our search for loopholes and loose ends in the Euclidean approach, we finally hit on the crucial idea, the one ingredient absolutely necessary to make the stir fry come out right: the universe must encode what physicists call causality. Causality means that empty space-time has a structure that allows us to distinguish unambiguously between cause and effect. It is an integral part of the classical theories of special and GTR.
>
> Euclidean quantum gravity does not build in a notion of causality. The term 'Euclidean' indicates that space and time are treated equally. The universes that enter the Euclidean superposition have four spatial directions instead of the usual one of time and three of space. Because Euclidean universes have no distinct notion of

> time, they have no structure to put events into a specific order; people living in these universes would not have the words 'cause' or 'effect' in their vocabulary. Hawking and others taking this approach have said that 'time is imaginary,' in both a mathematical sense and a colloquial one. Their hope was that causality would emerge as a large-scale property from microscopic quantum fluctuations that individually carry no imprint of a causal structure. But the computer simulations dashed that hope. (2008, p. 3)

In their Causal Dynamical Triangulations approach, causality is demonstrated to be a necessary *fundamental* feature for a cosmogony yielding our observable universe. Lee Smolin explains the implication:

> Some of the most widely believed ideas about quantum gravity are in fact wrong. For example, Stephen Hawking and others used to argue that causal structure was inessential, and that calculations could be done in quantum gravity by ignoring the differences between time and space – differences that exist even in relativity theory – and treating time as if it were another dimension of space (imaginary time) . . . . Ambjorn and Lolle's results show that this idea is wrong . . ..
>
> In particular it was shown . . . that if no restriction respecting causality is put in, no classical space-time geometry emerges . . ..
>
> One of the rules that Loll and Ambjorn impose is that each quantum space-time has to be seen as a sequence of possible spaces that succeed one another, like the ticks of a universal clock. The time coordinate, it is argued, is arbitrary, as in GTR, but the fact that the history of the world can be seen as a succession of geometries that succeed one another is not. (2006, pp. 242–3)

Interestingly, the trajectory of quantum gravity research is tending to correct the 'imaginary time' approach and in so doing to affirm the importance of consistent *earlier than/later than* relations and fundamental causality.

4. The 'initial' null topology is misconceived. J. Richard Gott and Li Xin Li have criticized Vilenkin (and Hartle–Hawking's) creation *ex nihilo* approach on two grounds. First, transitions in QM are always between allowed classical states. But Vilenkin and Hartle–Hawking's approach has a transition from a classically forbidden region to a classically allowed region. Second, the Vilenkin and Hartle–Hawking approaches should contain realistic energy fields (something closer

to what we actually see in nature). If they did, then Heisenberg's uncertainty principle would require that the initial state of their models have a finite and nonzero energy. If that is the case, then semi-classical quantum models like Vilenkin's and Hartle–Hawking's actually start in a classically allowed, metastable state, rather than 'nothing'. Gott and Li elaborate:

> The problem with this model (Vilenkin and Hartle–Hawking) is that it ignores the 'zero-point energy.' If there is a conformal scalar field ϕ, then the 'energy' levels should be $E_n = n + 1/2$. Even for $n = 0$ there is a 'zero-point-energy.' The potential makes the system behave like a harmonic oscillator in the potential well near $a = 0$. A harmonic oscillator cannot sit at the bottom of the potential well – the uncertainty principle would not allow it. There must be some zero-point-energy and the particle must have some momentum, as it oscillates within the potential well when the field ϕ is included. Thus, when the 'zero point-energy' is considered, we see that the initial state is not a point but a tiny oscillating ($0 < a < a_1$) big bang universe, that oscillates between big bangs and big crunches (though the singularities at the big bangs and big crunches might be smeared by quantum effects). This is the initial *classical* state from which the tunneling occurs. *It is metastable, so this oscillating universe could not have existed forever: after a finite half-life, it is likely to decay.* It reaches maximum radius a_1, and then tunnels to a classical de Sitter state at minimum radius a_2 where $a_2 < a_0$. (1998, p. 38; our emphasis)

Thus, we seem to have the same sort of situation that we encountered with respect to the Emergent model with its associated metastable ESS. The universe cannot be past eternal because the initial metastable state can have had only a finite lifetime. The Gott–Li interpretation seems to be a reasonable option for a realist interpretation of these models.[32] It employs known, meaningful interpretations of physical phenomena from 'classical' quantum theory and extends them to the quantum gravity models. One avoids the problems associated with the novelty of asserting a zero-energy condition for the initial state (denied by the Heisenberg uncertainty principle), the novelty of asserting a quantum transition from a forbidden to a classically allowed state (normal quantum theory includes only transitions over or through forbidden regions from one allowed state to another), and it is consistent with more realistic energy fields. This alternative is, of course, also consistent with the second premise of the *kalam* cosmological argument.

6. Concluding Remarks

In sum, we have seen no good reason to deny that space-times featuring an initial cosmological singularity have a beginning; but even more fundamentally we have seen that the claim that the universe began to exist does not depend upon the singular nature of space-time. Moreover, our survey of contemporary cosmological thinking reveals significant evidence that non-singular models may themselves to be incompatible with an infinite past, and therefore singularity-free quantum gravitational approaches can be quite supportive of the *kalam* cosmological argument's second premise. So when Pitts worries, 'It is not implausible that some singularity-wielding theistic apologists will be tempted to resist scientific progress in the form of a new quantum theory of gravity in order to maintain an apologetic strategy in which they have invested' (2008, p. 19), one cannot help but wonder who is the object of so uncharitable an allegation. For we have seen that the second premise of the *kalam* cosmological argument enjoys considerable empirical support even if space-time is non-singular.[33]

Notes

1. Writing in the *Cambridge Companion to Atheism*, Quentin Smith reports, 'a count of the articles in the philosophy journals shows that more articles have been published about . . . the Kalam argument than have been published about any other philosopher's contemporary formulation of an argument for God's existence' (2007, p. 183). For the history of the argument, see Craig, 1980; for literature and a recent defence of the argument, see Craig and Sinclair, 2009.
2. In places Pitts recognizes this: 'The truth of the second premise [of the *kalam* cosmological argument], or rather, the source of warrant for the second premise if it is true, is the key question' (2008, p. 687; cf. p. 676).
3. So to argue (while avoiding the fallacy of composition) would be to challenge the truth, not of (2), but of (1) of the *kalam* cosmological argument. Sometimes Pitts does conflate what are really challenges to (1) with putative objections to (2). For example, his mention of four-dimensionalism as rendering space-time 'self-explanatory' (2008, p. 688) is a case in point. Again, Pitts's Cosmic Destroyer argument is really an objection to (1): if the coming-into-being of the universe must have a cause, then so must its ceasing-to-be.
4. One cannot, in passing, help but wonder how imposing topological constraints on the doctrine of creation in time, as Pitts would do, is modally enriching, as he claims, since then there are possible worlds (like, apparently, the actual world) in which the universe lacks a first moment and so cannot be said on Pitts's account to have been created in time by God.

5. It is noteworthy that Earman, whose objection Pitts echoes, does not in the end deny the acceptability of our explication of 'begins to exist' but merely advises that on such a reading premise (1) 'is not an obvious "metaphysical truth"'; in particular (1) does not follow from the principle that *every event has a cause*, which is satisfied in general relativistic cosmology (1995, p. 208).
6. NB that both Pitts's argument and our response are limited to a faux universe where GTR is the fundamental theory of gravity.
7. We also note that there is no good reason to think that metric conventionalism holds in the actual world. As has been argued elsewhere (Craig, 2001, ch. 2), metric conventionalism is an implausible thesis for which no good arguments exist.
8. It is at least worth noting that Pitts, not content with the promise of current quantum cosmology, goes so far as to champion the rights of *unborn* and, indeed, *unconceived* hypotheses as grounds for scepticism about (9): 'The relevant set of competitors for GTR includes the set of theories that agree with GTR on all experiments to date, whether already entertained on Earth or not. This set might be infinite, might well be large, likely contains several members, and almost certainly has at least one member, a quantum theory of gravity. The set most likely has at least one member that resolves the singularities of GTR. Thus, in the strong field regime it is not at all clear why one should take GTR seriously' (2008, p. 696). Pitts's final sentence is a non sequitur. As Timothy McGrew has remarked in personal conversation, in order to mount a significant challenge to GTR it is not enough to show that it is probable that an unknown theory of quantum gravity exists which shares GTR's empirical adequacy to date and resolves the singularities of GTR. Pitts also needs to show that this unknown competing theory has, by comparison to GTR, a non-negligible prior probability. For there may be other overriding factors that are relevant to its probability, for example, its want of simplicity or elegance. Since these theories are unconceived, it is hard to see how Pitts can know such a thing. In any case, as we remark in the text, the important issue is not whether that unknown theory resolves the singularities of GTR but whether the physical universe as described by the theory is infinitely extrapolable to the past. What one wants from Pitts is some evidence or argument that there is an unknown theory of quantum gravity meeting all the desiderata.
9. Our intention is to provide a complementary update to the discussion in Craig and Sinclair, 2009. Thus, for example, the material in the original essay had significant content regarding string-based cosmogonies which will not be repeated here.
10. It also permits us to regard records of the past, such as our own memories, as being relics of a true reality as opposed to illusory. We will hold that starlight that appears in our telescopes in the present actually came from a distant star in a real past. So called 'top–down' interpretations of the Feynman path integral approach to quantum gravity (Hawking, 2003), or 'decoherent histories' (see Hartle, 1998; Dowker and Kent, 1996), are examples of quantum gravity approaches that would deny the objective reality of a unique past that *evolved* into the present. This doesn't mean, necessarily that these approaches must be rejected. Rather, we reject the *interpretation* of the formalism. Hawking, in

the above reference, says about his own theory that 'One can interpret this in the bottom up picture as the spontaneous creation of an inflating universe from nothing.' On this view, we have precisely the creation *ex nihilo* that lends evidence towards *kalam*'s second premise.

11. This implication was affirmed by Vilenkin in a personal communication on 4 March 2004. Note that this conclusion would apply to the well-known Ekpyrotic cosmogony of Paul Steinhardt and Neil Turok mentioned by Monton. Their model has a genuine past boundary and thus accords with the second premise of the *kalam* cosmological argument. See Craig and Sinclair, 2009 for a more detailed explanation.
12. See also the collection of papers by Brett McInnis (2007a, 2007b, 2008). McInnis provides an alternative argument for why inflation must have an ultimate origin. He suggests that in string theory (for purely geometric reasons) the arrow of time can only originate in a topological creation event 'from nothing' such as in the Vilenkin proposal (see section 5 of our paper). Subsequent bubble universes, whether embedded or distinct disconnected space-times (see Carroll and Chen, 2004), will have an arrow of time only if they *inherit* it from their mother. Thus, given our observed arrow of time, we know we have an ultimate 'mother' that came into being via creation *ex nihilo*.
13. To illustrate, one of the authors once participated in a similar exercise while working in the defence industry. The author noted that a particular air-to-air combat model explicitly included missile flyouts against air targets. A much more elegant (and efficient) model was built by deriving a set of equations whereby time and distance intervals could be dropped from the formalism. The combat could be described by an impact sequence of missile salvos that does not reference the time or space intervals between salvo arrivals. Now would the mere fact of the existence of this model (and its superiority in terms of simplicity, explanatory scope etc.) imply that time and space intervals do not exist? This would be a ludicrous conclusion, as any experience as a combat aviator would demonstrate. Air superiority is maintained by having the superior range missile (a distance interval), or shortest flight (time interval) such that one's missile gets there *first*. It is precisely the intervals that produce salvo impact order and hence air superiority. Individual life and death, even the rise and fall of empires, ride on this conclusion. One is reminded of Einstein's advice that we should make our models as simple as possible, but not *too* simple.
14. See Aguirre, 2007 for a detailed attempt to circumvent BGV in favour of a time-reversed or 'emergent' non-singular cosmogony. These suggestions of Aguirre's are considered in sections 4.4 and 4.6.
15. See, e.g., Earman and Mosterin, 1999, for a related argument.
16. See also Damour and Henneaux (2000, p. 923), whose results update B–K–L to consider quantum string cosmogonies: 'our findings suggest that the spatial inhomogeneity continuously increases toward a singularity, as all quasi-uniform patches of space get broken up into smaller and smaller ones by the chaotic oscillatory evolution. In other words, the space-time structure tends to develop a kind of "turbulence".'
17. Ellis does indicate that a finite past is an acceptable possible form for an Emergent model (Ellis and Maartens, 2004, section V, p. 5).

18. We take as real the scale factor behaviour, while seeing the function of scalar field vs. time as instrumental. Otherwise, the scalar field could itself perform the function of a clock.
19. The quantum state must not be infinitely squeezed to be semi-classical (see section 4.5); the centre equilibrium point only exists if the semi-classical assumption holds.
20. See Craig and Sinclair (2009) for an extensive discussion of another asymptotically static cosmogony, the string-based pre-Big Bang Inflation model of Maurizio Gasperini and Gabrielle Veneziano. See Kaloper et al. (1998) for a similar criticism based on fluctuation of the dilaton field.
21. Also see Tavakol and Carneiro (2009). The argument is that the past is infinite. But a suggested method of leaving the stabilized ESS so as to allow an inflationary expansion invokes a quantum tunnelling from a local minimum in potential (stabilized ESS) to a global minimum (open-ended inflation). But this local minimum is a metastable state with a finite lifetime. Thus the model is not past infinite.
22. Examples in the recent literature of asymptotically static non-singular cosmogonies include, e.g., Falciano and Pinto-Neto, 2009; Seahra and Böhmer, 2009; and Barrow and Tsagas, 2009. The first compares the predictions of such models with quantum scalar perturbations observed in the cosmic background radiation via the Wilkinson Microwave Anisotropy Probe and casts doubt on their ultimate viability. The second and third investigate the viability of a past eternal ESS model. The Seahra and Böhmer paper shows that an ESS is unstable to some type of perturbation for generic quantum gravity theories and generic equation of state (that is, the type of matter/energy residing in the universe) for f(R) gravity theories. The Barrow and Tsagas paper exploits a stability region for ESS for 'phantom' or 'ghost' matter universes. See Craig and Sinclair, 2009 for a discussion of some of these 'phantom bounce' cosmogonies.
23. Banks complains, 'I have a problem with ALL cyclic cosmogonies.... The collapsing phase of these models always has a time dependent Hamiltonian for the quantum field fluctuations around the classical background. Furthermore the classical backgrounds are becoming singular. This means that the field theories will be excited to higher and higher energy states (define energy in some adiabatic fashion during the era when the cosmology is still fairly slowly varying, and use this to classify the states, even though it is not conserved). High energy states in field theory have the ergodic property – they thermalize rapidly, in the sense that the system explores all of its states. Willy Fischler and I proposed that in this situation you would again tend to maximize the entropy. We called this a Black Crunch and suggested the equation of state of matter would again tend toward $p = \rho$. It seems silly to imagine that, even if this is followed by a re-expansion, that one would start that expansion with a low entropy initial state, or that one had any control over the initial state at all' (pers. comm., 12 October 2008). An important caveat is that a general description of entropy in the presence of gravity (either in GTR or the still undiscovered Quantum Gravity) has never been achieved. But there are some special cases that are thought reliable by most cosmologists (see Carroll et al., 2007, slides 11 and 12), such as the entropy of a thermal gas, black holes and cosmic horizons (this last being related to a

positive cosmological constant). These three cases are thought to represent the dominant entropy contributors to the universe for three different eras: the early universe (thermal gas), the present (black holes) and the far future (a 'pure' de Sitter-like cosmic horizon). If these special cases are accurate, they demonstrate that entropy increases as the universe ages, thus providing some proof of a generalized second law of thermodynamics.

24. See also Coule (2008), where a suggestion is made that maximum entropy is capped by the holographic bound (cosmic horizon entropy for a de Sitter universe), and that the entropy of a bounce violates this restriction.
25. A philosopher sympathetic with a tensed theory of time would have an additional rationale for believing that such models really represent a multiverse propagating from a common origin. There would be a reason why the cosmological, thermodynamic, electromagnetic and psychological arrows of time would all align; they are physical manifestations of the same underlying temporal becoming of metaphysical time.
26. Note that Smith and Pitts are interested in different topological features of space-time. Pitts's concern is whether the universe has a first moment. Smith's concern is whether the topology of a space-time manifold (e.g. unbroken sheet, torus, figure eight etc.) is more fundamental than its metrical aspects.
27. '[T]he question of whether the universe had a beginning at a finite time is now "transcended". At first, the answer seems to be "no" in the sense that the quantum evolution does not stop at the Big Bang. However, since space-time geometry "dissolves" near the big-bang, *there is no longer a notion of time, or of "before" or "after" in the familiar sense*. Therefore, strictly, the question is no longer meaningful. The paradigm has changed and meaningful questions must now be phrased differently, without using notions tied to classical space-times' (Ashtekar, Bojowald and Lewandowski, 2003, p. 263; our emphasis).
28. 'And why would we not apply Occam's razor to the pre-history?' (Bojowald, 2009, p. 16)
29. A creation *ex nihilo* makes more sense than positing time's emergence 'from' a pure quantum state. Augustine seems to have had a similar insight 1600 years ago: 'We can correctly say, "There was a time when Rome did not exist: there was a time when Jerusalem, or Abraham, or man, or anything of this kind did not exist". We can in fact say, "There was a time when the world did not exist", if it is true that the world was created not at the beginning of time, but some time after. But to say, "There was a time when time did not exist", is as nonsensical as to say, "There was a man when no man existed", or, "This world existed when the world was not" ' (1984, *City of God*, Bk XII, ch. 16).
30. Also see McInnis (2007a, 2007b, 2008) for a string-based model of this type.
31. 'Creation of a universe from nothing . . . is a transition from the null topological sector containing no universes at all to the sector with one universe of topology S^3' (Vilenkin, 1994, p. 23).
32. In Craig and Sinclair, 2009, we profile Gott and Li's own attempt at solving cosmology's origins problem based on the postulate of closed timelike curves (a time machine). Unfortunately, they can avoid Stephen Hawking's Chronology Protection Conjecture only by asserting an 'initial' cosmic

vacuum which is measure zero with respect to all possible configurations. In other words, their model is infinitely fine-tuned.

33. We thank Martin Bojowald, Thomas Banks, George Ellis, Alexander Vilenkin, Donald Page, Quentin Smith, Timothy McGrew, Rüdiger Vaas and Christian Böhmer for useful discussions and comments.

References

Aguirre, Anthony (2007) 'Eternal Inflation Past and Future', in R. Vaas (ed.), *Beyond the Big Bang* (Heidelberg: Springer Verlag); preprint: http://arxiv.org/abs/0712.0571.

Aguirre, Anthony, and Steve Gratton (2002) 'Steady State Eternal Inflation', *Phys. Rev. D* 65, 083507; preprint: http://arxiv.org/abs/astro-ph/0111191v2.

Al-Ghazali (1962) *Kitab al-Iqtisad fi'l-Iqtiqad* (Ankara: University of Ankara Press).

Al-Ghazali (1963) *Tahafut al-Falasifah*. Trans. S. A. Kamali (Lahore: Pakistan Philosophical Congress).

Ashtekar, A., M. Bojowald and J. Lewandowski (2003) 'Mathematical Structure of Loop Quantum Cosmology', *Adv.Theor.Math.Phys*. 7: 233–68; preprint: http://arxiv.org/abs/gr-qc/0304074.

Ashtekar, A., Tomasz Pawlowski and Parampreet Singh (2006) 'Quantum Nature of the Big Bang', *Phys. Rev. D* 74, 084003; preprint: http://arxiv.org/abs/gr-qc/0602086v2.

Augustine (1984) *City of God* (London: Penguin Books).

Asimov, Isaac (1989) 'The Relativity of Wrong', *The Skeptical Inquirer* 14.1 (Fall); available online at http://chem.tufts.edu/AnswersInScience/RelativityofWrong.htm.

Banks, T., and W. Fischler (2002) 'Black Crunch'; preprint: http://arxiv.org/abs/hep-th/0212113v1.

Barrow, J., and M. Dabrowski (1995) 'Oscillating Universes', *Monthly Notices of the Royal Astronomical Society* 275: 850–62.

Barrow, J., and C. Tsagas (2009) 'On the Stability of Static Ghost Cosmologies'; available online at http://arxiv.org/abs/0904.1340v1 (21 May 2009).

Belinsky, V. A., I. M. Khalatnikov and E. M. Lifshitz (1970) 'Oscillatory Approach to a Singular Point in the Relativistic Cosmology', *Advances in Physics* 19: 525–73.

Bojowald, M. (2005) 'Original Questions', *Nature* 436: 920–1.

Bojowald, M. (2006) 'Universe Scenarios from Loop Quantum Gravity', *Annalen Phys*. 15: 326–41; preprint: http://arxiv.org/abs/astro-ph/0511557v1.

Bojowald, M. (2007) 'Large Scale Effective Theory for Cosmological Bounces', *Phys. Rev. D* 74, 081301; preprint: http://arxiv.org/abs/gr-qc/0608100v2.

Bojowald, M. (2009) 'A Momentous Arrow of Time', in L. Mersini-Houghton (ed.), *The Arrow of Time* (Heidelberg: Springer Verlag); preprint http://arxiv.org/abs/0910.3200.

Bojowald, M., R. Maartens and P. Singh (2004) 'Loop Quantum Gravity and the Cyclic Universe', *Phys. Rev. D* 70, 083517; preprint: http://arxiv.org/abs/hep-th/0407115.

Bojowald, M., G. Date and G. Hossain (2004) 'The Bianchi IX Model in Loop Quantum Cosmology', *Class. Quant. Grav.* 21, 3541; preprint: http://arxiv.org/abs/gr-qc/0404039.

Bojowald, M., and R. Tavakol (2008a) 'Recollapsing Quantum Cosmologies and the Question of Entropy', *Phys. Rev. D* 78, 023515; preprint: http://arxiv.org/abs/0803.4484v1.

Bojowald, M., and R. Tavakol (2008b) 'Loop Quantum Cosmology: Effective Theories and Oscillating Universes', in R. Vaas (ed.), *Beyond the Big Bang* (Heidelberg: Springer Verlag); preprint: http://arxiv.org/abs/0802.4274v1.

Borde, A., A. Guth and A. Vilenkin (2003) 'Inflationary Spacetimes Are not Past-Complete', *Physical Review Letters* 90, 151301; preprint: http://arxiv.org/abs/gr-qc/0110012.

Carroll, Sean (2008) 'What If Time Really Exists?', FXQI essay contest on the nature of time, available online at www.fqxi.org/community/forum/topic/318.

Carroll, Sean et al. (2007) 'Why Is the Past Different from the Future: The Origin of the Universe and the Arrow of Time', available online at http://preposterousuniverse.com/talks/time-colloq-07/.

Carroll, Sean, and Jennifer Chen (2004) 'Spontaneous Inflation and the Origin of the Arrow of Time'; preprint: http://arxiv.org/abs/hep-th/0410270v1.

Coleman, S., and F. De Luccia (1980) 'Gravitational Effects on and of Vacuum Decay', *Phys. Rev. D* 21, 3305.

Copan, P., and W. L. Craig (2004) *Creation out of Nothing* (Grand Rapids, MI: Baker).

Coule, D. H. (2008), 'Holography Constrains Quantum Bounce'; preprint: http://arxiv.org/abs/0802.1867.

Craig, W. L. (2006) 'Naturalism and Cosmology', in A. Corradini, S. Galvan and E. J. Lowe (eds), *Analytic Philosophy without Naturalism* (Routledge Studies in Contemporary Philosophy; New York: Routledge), pp. 97–133.

Craig, W. L. (1980) *The Cosmological Argument from Plato to Leibniz* (London: Macmillan & Co.).

Craig, W. L. (2001) *Time and the Metaphysics of Relativity* (Philosophical Studies Series, 84; Dordrecht: Kluwer Academic Publishers).

Craig, W. L., and J. Sinclair (2009) 'The *Kalam* Cosmological Argument', in W. L. Craig and J. P. Moreland (eds), *Blackwell Companion to Natural Theology* (Oxford: Blackwell), pp. 101–201.

Damour, T., and M. Henneaux (2000) 'Chaos in Superstring Cosmology', *Physical Review Letters* 85: 920–3; preprint: http://arxiv.org/abs/hep-th/0003139.

Dowker, F., and A. Kent (1996) 'On the Consistent Histories Approach to Quantum Mechanics', *J. Statist. Phys.* 82: 1575–1646; preprint: http://arxiv.org/abs/gr-qc/9412067v2.

Earman, J. (1995) *Bangs, Crunches, Whimpers, and Shrieks: Singularities and Acausalities in Relativistic Spacetimes* (New York: Oxford University Press).

Earman, J., and J. Mosterin (1999) 'A Critical Look at Inflationary Cosmology', *Philosophy of Science* 66: 1–49.

Ellis, G. F. R., and R. Maartens (2004) 'The Emergent Universe: Inflationary Cosmology with No Singularity', *Classical and Quantum Gravity* 21: 223; preprint: http://arxiv.org/abs/gr-qc/0211082.

Ellis, G. F. R., Jeff Murugan and Christos Tsagas (2004) 'The Emergent Universe: An Explicit Construction', *Class. Quant. Grav.* 21: 233–50; preprint: http://arxiv.org/abs/gr-qc/0307112.

Falciano, F. T., and N. Pinto-Neto (2009) 'Scalar Perturbations in Scalar Field Quantum Cosmology', *Phys. Rev. D* 79, 023507; preprint: http://arxiv.org/abs/0810.3542.

Geroch, R. (1978) *GR from A to B* (Chicago and London: University of Chicago Press).

Gott, J. R. III, and L.-X. Li (1998) 'Can the Universe Create Itself?', *Phys. Rev. D* 58: 2, 023501–1.

Hartle, J. (1998) 'Quantum Pasts and the Utility of History', *Phys. Scripta* T76: 67; preprint: http://arxiv.org/abs/gr-qc/9712001v1.

Hartle, J., and S. Hawking (1983) 'The Wave Function of the Universe', *Phys. Rev. D* 28: 12, 2960–75.

Hawking, S. (2003) 'Cosmology from the Top Down', talk presented at the Davis Inflation Meeting; preprint: http://arxiv.org/abs/astro-ph/0305562.

Hawking, S., and R. Penrose (1970) 'The Singularities of Gravitational Collapse and Cosmology', *Proceedings of the Royal Society of London A* 314: 529–48.

Hawking, S., and R. Penrose (1996) *The Nature of Space and Time* (Princeton, NJ: Princeton University Press).

Jurkiewicz, J., L. Loll and J. Ambjorn (2008) 'Using Causality to Solve the Puzzle of Quantum Spacetime', *Scientific American* (July); available online at www.scientificamerican.com/article.cfm?id=the-self-organizing-quantum-universe.

Kaloper, N., A. Linde and R. Bousso (1998) 'Pre-Big-Bang Requires the Universe to Be Exponentially Large from the Very Beginning', *Phys. Rev. D* 59, 043508; available online at http://hep-th/9801073.

Kiefer, Claus (2008) 'Does Time Exist in Quantum Gravity?', FXQI essay contest on the Nature of Time; available online at www.fqxi.org/data/essay-contest-files/Kiefer_fqx.pdf.

Kiefer, Claus (2009) 'Can the Arrow of Time Be Understood from Quantum Cosmology?', in L. Mersini-Houghton (ed.), *The Arrow of Time* (Heidelberg: Springer Verlag); preprint: http://arxiv.org/abs/0910.5836v1.

Kiefer, C., and H. D. Zeh (1995) 'Arrow of Time in a Recollapsing Quantum Universe', *Phys. Rev. D* 51: 4145–53; preprint: http://arxiv.org/abs/gr-qc/9402036v2.

Loll, R. (2008) 'The Emergence of Spacetime, or, Quantum Gravity on Your Desktop', *Classical Quantum Gravity* 25.11: 114006; available online at http://arxiv.org/abs/0711.0273v2.

Mersini-Houghton, Laura (2009) 'Notes on Time's Enigma', FXQI conference on time, Azores; available online at http://arxiv.org/abs/0909.2330v1.

McInnis, Brett (2007a) 'Arrow of Time in String Theory', *Nucl. Phys. B* 782: 1–25; preprint: http://arxiv.org/abs/hep-th/0611088v3.

McInnis, Brett (2007b) 'The Arrow of Time in the Landscape', in R. Vaas (ed.), *Beyond the Big Bang* (Springer: Heidelberg); preprint: http://arxiv.org/abs/0711.1656v2.

McInnis, Brett (2008) 'Initial Conditions for Bubble Universes', *Phys. Rev. D* 77, 123530; preprint: http://arxiv.org/abs/0705.4141v5.

Monton, Bradley (forthcoming) 'Prolegomena to Any Future Physics-Based Metaphysics', *Oxford Studies in Philosophy of Religion* 3.

Mulryne, David, Reza Tavakol, James E. Lidsey and George F. R. Ellis (2005) 'An Emergent Universe from a Loop', *Phys. Rev. D* 71, 123512; preprint: http://arxiv.org/abs/astro-ph/0502589v1.

Oppy, G. (2006) *Arguing about Gods* (Cambridge: Cambridge University Press).

Overbye, D. (2006) '9 billion-year-old "dark energy" reported', *The New York Times* (15 January).

Parikh, M., Ivo Savonije and Erik Verlinde (2003) 'Elliptic de Sitter Space', *Phys. Rev. D* 67, 064005; preprint: http://arxiv.org/abs/hep-th/0209120v2.

Penrose, R. (2005) *The Road to Reality* (New York: Alfred A. Knopf).

Penrose, R. (2006) 'Before the Big Bang: An Outrageous New Perspective and Its Implications for Particle Physics', in *Proceedings of the European Particle Accelerator Conference (EPAC) 2006* (Edinburgh: EPS–AG), pp. 2759–62; available online at http://accelconf.web.cern.ch/AccelConf/e06/PAPERS/THESPA01.PDF.

Penrose, R. (2009) 'Black Holes, Quantum Theory and Cosmology', in *J. Phys.: Conf. Ser.* 174, 01; available online at http://iopscience.iop.org/1742-6596/174/1/012001/pdf/1742-6596_174_1_012001.pdf.

Pitts, J. B. (2008) 'Why the Big Bang Singularity Does Not Help the Kalam Cosmological Argument for Theism', *British Journal for the Philosophy of Science* 59.4: 675–708.

Quine, W. V. (1969) 'Epistemology Naturalized', in idem, *Ontological Relativity and Other Essays* (New York: Columbia University Press), pp. 69–90.

Rea, Michael C. (2002) *World without Design: The Ontological Consequences of Naturalism* (Oxford: Clarendon Press).

Seahra, Sanjeev S., and C. G. Böhmer (2009) 'Einstein Static Universes Are Unstable in Generic f(R) Models', *Phys. Rev. D* 79, 064009; preprint: http://arxiv.org/abs/0901.0892.

Smith, Q. (1985) 'On the Beginning of Time', *Noûs* 19: 579–84.

Smith, Q. (1993) 'The Wave Function of a Godless Universe', in Quentin Smith and William Lane Craig (eds), *Theism, Atheism and Big Bang Cosmology* (Oxford: Clarendon Press, 1993), pp. 301–37.

Smith, Q. (1997) 'The Ontological Interpretation of the Wave Function of the Universe', *The Monist* 80.1: 160–85.

Smith, Q. (2000) 'The Black Hole Origin Theory of the Universe: Frontiers of Speculative, Current Physical Cosmology'; available online at www.faculty.umb.edu/gary_zabel/Courses/Parallel%20Universes/Texts/the_black_hole_origin_theory_of_the_universe_frontiers_of_s.htm.

Smith, Q. (2007) 'Kalam Cosmological Arguments for Atheism', in M. Martin (ed.), *The Cambridge Companion to Atheism* (Cambridge Companions to Philosophy; Cambridge: Cambridge University Press), pp. 182–98.

Sorabji, R. (1983) *Time, Creation, and the Continuum* (Ithaca, NY: Cornell University Press).

Sorkin, R. D., and D. P. Rideout (1999) 'Classical Sequential Growth Dynamics for Causal Sets', *Phys. Rev. D* 61, 024002; available online at http://arxiv.org/abs/gr-qc/9904062v3.

Smolin, Lee (2006) *The Trouble with Physics* (New York: Houghton Mifflin).

Smolin, Lee (2009) 'The Unique Universe', *PhysicsWorld* (June); available online at http://physicsworld.com/cws/article/print/39306.

Tavakol, R., and S. Carneiro (2009) 'Stability of the Einstein Static Universe in the Presence of Vacuum Energy', *Phys. Rev. D* 80, 043528; preprint: http://arxiv.org/abs/0907.4795v2.

't Hooft, Gerard (2009) 'Quantum Gravity without Space-time Singularities or Horizons', presented at the Eric Summerschool of Subnuclear Physics, available online at http://arxiv.org/abs/0909.3426.

Tod, K. P. (2003) 'Isotropic Cosmological Singularities: Other Matter Models', *Class. Quantum Gravity* 20: 521–34.

Tolman, R. C. (1934) *Relativity, Thermodynamics and Cosmology* (Oxford: Clarendon Press).

Vaas, R. (2004) 'Time before Time: Classifications of Universes in Contemporary Cosmology, and How to Avoid the Antinomy of the Beginning and Eternity of the World', in W. Löffler and P. Weingartner (eds), *Knowledge and Belief. Papers of the 26th International Wittgenstein Symposium* (Kirchberg am Wechsel: Austrian Ludwig Wittgenstein Society), pp. 351–3; preprint: http://arXiv.org/abs/physics/0408111.

Veneziano, G., and M. Gasperini (2002) 'The Pre Big Bang Scenario in String Cosmology', *Physics Reports* 373: 1; preprint: http://arxiv.org/abs/hep-th/0207130.

Vilenkin, A. (1982) 'Creation of Universes from Nothing', *Phys. Lett.* 117B: 25–28.

Vilenkin, A. (1994) 'Approaches to Quantum Cosmology', *Phys. Rev. D* 50: 2581–94; preprint: http://lanl.arxiv.org/abs/gr-qc/9403010v1.

Vilenkin, A. (2002) 'Quantum Cosmology and Eternal Inflation', in *The Future of Theoretical Physics and Cosmology: Proceedings of the Conference in Honor of Stephen Hawking's 60th Birthday*; preprint: http://arxiv.org/abs/gr-qc/0204061.

Vilenkin, A. (2006) *Many Worlds in One* (New York: Hill & Wang).

Zeh, H. D. (2009) 'Open Questions Regarding the Arrow of Time', in L. Mersini-Houghton (ed.), *The Arrow of Time* (Heidelberg: Springer Verlag); preprint http://arxiv.org/abs/0908.3780.

6
The Theistic Multiverse: Problems and Prospects

Klaas J. Kraay

In recent decades, there has been astonishing growth in scientific theorizing about multiverses. Once considered outré or absurd, multiple universe theories appear to be gaining considerable scientific respectability. There are, of course, many such theories, including (1) Everett's (1957) *many worlds interpretation* of quantum mechanics, defended by Deutsch (1997) and others; (2) Linde's (1986) *eternal inflation* view, which suggests that universes form like bubbles in a chaotically inflating sea; (3) Smolin's (1997) *fecund universe* theory, which proposes that universes are generated through black holes; (4) the *cyclic model*, recently defended using string/M theory by Steinhardt and Turok (2007), which holds that distinct universes are formed in a never-ending sequence of Big Bangs and Big Crunches; and (5) Tegmark's (2007) 'Level IV' multiverse, which contains many universes governed by distinct mathematical and scientific laws. While not all of these preclude each other, the details and implications of each one are hotly contested.[1]

In one area within the philosophy of religion (the debate concerning the 'fine-tuning' argument), scientific multiverse theories are widely held to be hostile to theism. This is because such theories appear to account for the relevant data – the biophilic parameters of the universe we inhabit – without appeal to an intelligent designer. Yet, in recent years, several philosophers[2] and one physicist[3] have offered reasons for thinking that if theism is true, the actual world comprises (or probably comprises) many universes. I first set out some requirements – both scientific and otherwise – for such a theory. I then survey some problems such theories are held to face, and some prospects they are thought to have. Finally, I examine arguments both for and against the claim that multiverse theories can help theists respond to the problem of evil. I conclude that such theories advantage neither the theist nor the

atheist in the debate about evil: they merely require *reframing* arguments from evil.

1. Some Requirements for a Theistic Multiverse Theory

The claim that on theism the actual world is (or probably is) a multiverse is not explicitly about God's activity with respect to the natural world. But of course theism holds that God is the ultimate causal explanation for the existence of contingent reality, including the natural realm we inhabit. Such a view must, then, be compatible with what contemporary science has to say about causation. Theists typically think of God's causal role as an initial act of creation *ex nihilo*, followed by the continued conservation in existence of all the things there are. This requires an account of causation consistent with the doctrine of an immaterial being bringing about physical effects, which can allow for such a being to bring about effects *ex nihilo*. There are related issues about God and time: for example, if God is thought to exist outside of time, an account is needed which can allow for a timeless being to have temporal effects. A model of causation is needed which can make sense of the traditional idea that God *sustains* things in existence. And famously, of course, there are vexing questions about how best to understand traditional claims about God's causal interventions in the course of nature. In what follows, however, I set all these issues aside, since they pertain to many theistic accounts of the existence of contingent reality, and are not peculiar to multiverse theories.

A complete theistic multiverse theory will offer an account, consistent with the best science, of what a *universe* is.[4] But, of course, scientific research has not yet fully answered this question: apart from some very general points of consensus, different theories vary widely.[5] Theistic multiverse theories will not, then, be complete until the science is settled. But in the meantime, defenders of theistic multiverses can develop their views in accordance with one or more scientific theories, recognizing that these are controversial, or they can remain neutral between many or all such theories. Most theistic multiverse theories take for granted that universes are independent, non-interacting, spatiotemporally interrelated objects. Some, however, appear to assume that all universes are related, either temporally (Stewart, 1993, p. 61), or by being embedded in a higher spatial dimension (Hudson, 2006). Some philosophers are officially neutral on this issue (Draper, 2004, p. 316; Forrest, 1996, p. 218). I take it that whether or not there are completely disconnected space-times is a question for science, not theology

or philosophy, to settle. It's worth noting that if there are, this will falsify philosophical theories, like David Lewis's *modal realism*, which preclude this.[6] And this matters in the current context because at least one *rival* to theistic multiverse theories depends upon Lewis's view that there cannot be independent space-times.[7] So here is a clear case where science can help to adjudicate between competing philosophical theories.

Some defenders of a theistic multiverse maintain that universes come in different *kinds* or *types*, without giving a complete account of what they mean. Stewart appears to think that universes are to be distinguished with respect to the divine 'values' and 'purposes' that they express, but offers no details (1993, p. 61). O'Connor speaks of 'organic kinds' and 'value types', without definitions or examples (2008, p. 120).[8] Forrest (1996, p. 116) suggests that God will create 'universes of many types', without saying what a type is. In an earlier paper, he suggests that universes can be classified into 'kinds', according to the laws of nature they exhibit (1981, p. 50). For this suggestion to be defended, it must be shown compatible with scientific accounts of the laws of nature. More generally, any view that relies on such a classification scheme for universes must defend it, and this involves ensuring that it is consistent with the best contemporary science.

Some authors claim that there can be *duplicate* universes within a multiverse (Parfit, 1992, p. 423; Monton, forthcoming); but others deny this, citing the Principle of Identity of Indiscernibles (McHarry, 1978, p. 133; Turner, 2003, p. 150). Still others are silent or neutral on this issue (Stewart, 1993; Forrest, 1981, 1996; Draper, 2004; Hudson, 2006; O'Connor, 2008; Kraay, 2010a). Of course this principle is controversial, so those who maintain that there cannot be duplicate universes should either offer plausible arguments for it, or else should offer some other reason for thinking that there cannot be duplicates. But it is also worth noting that, even if the Principle of Identity of Indiscernibles is false, in principle there could be scientific reasons which preclude there being duplicate universes. So physics and cosmology will be relevant here as well.

Some authors appear attracted to the view that, on theism, *every* universe will feature living creatures (Leslie, 1989, p. 102; Turner, 2003, p. 147; O'Connor, 2008, p. 112). For Turner and O'Connor, this is because a loving God would want to show love to creatures. While God may want to have creatures to love, there may of course be *other* reasons why God would create uninhabited universes. In any case, this is another point at which scientific research places a constraint on theological speculation: if science establishes that there are uninhabited

universes, as all current scientific multiverse theories maintain, these claims will need to be abandoned.

Explicitly or implicitly, all theories of God's role as creator and sustainer involve God selecting, from an array of possible worlds, just one for actualization.[9] And it is generally thought that these objects of God's choice are proper subjects of axiological evaluation. So all theistic multiverse theories are committed to claims about the *axiological status* of both possible worlds and universes. Regarding the former, I have argued elsewhere that a good framework for this involves taking the axiological status of worlds to depend upon which world-good-making properties (WGMPs) and the world-bad-making properties (WBMPs) it exhibits, and, for degreed properties, the degree to which they are instantiated.[10] But of course this is just a schema: a complete theory would fill out the details of what these properties are, and just how they determine the axiological status of worlds. Ideally, such a completed theory will reveal whether there is a *best* or *unsurpassable* world (in various senses to be distinguished below), and whether or two or more worlds can be tied in axiological status.

Scientific research may constrain this axiological theorizing. Here are three examples of how this might happen. First, someone might hold that some or all axiological properties supervene on natural ones: if so, an account of the relevant natural properties is needed to ground axiological claims. Second, someone might hold with Leibniz that the best world features simple natural laws that generate great diversity: if so, an account of both laws and simplicity is needed. Third, someone might hold that worlds in which *everything* is spatiotemporally interconnected are, *ceteris paribus*, more valuable than those which feature disconnected space-times.[11] But if such interconnected worlds are precluded by science, then such claims about what is expectable on theism will have to be adjusted.

Suppose that axiological theorizing leads to a broad consensus about what sorts of properties are WGMPs and WBMPs. A further issue – which has received insufficient attention – is the *modal status* of these properties.[12] Most authors appear to assume *necessitarianism*: the view there is only one common set of WGMPs and WBMPs, with reference to which *all* possible worlds are to be evaluated. On the rival view, *contingentism*, there is no one common set of axiological properties relevant to all worlds: some worlds are properly evaluated with respect to one set of WGMPs and WBMPs, while other worlds are properly evaluated with respect to a distinct set of WGMPs and WBMPs. On necessitarianism, all possible worlds are *commensurable*: this is just what it means to say that

they are all to be evaluated with respect to the common set of axiological properties. On contingentism, however, there are failures of commensurability between possible worlds. On contingentism, it is useful to define a *world-cluster* to be a set of worlds which are commensurable. (Necessitarianism, then, amounts to view that there is just one world-cluster.) A complete axiological theory will settle whether necessitarianism or contingentism is true.[13]

On either view, however, there is a further complication. Although all worlds within a given cluster are *commensurable*, they may not all be *comparable*: there may be a pair of world-cluster-mates *w1* and *w2* such that neither is better than the other, nor are they *equal* in axiological status. If so, it will be useful to reserve the term *world-hierarchy* to refer to a set of worlds (within a given cluster) which are both commensurable *and* comparable. No two worlds belonging to different world-hierarchies are comparable, although they are commensurable only if they are cluster-mates. A complete axiological theory will settle whether there are genuine failures of comparability between commensurable worlds.

These distinctions are important because they bear on the question of whether or not there is a best possible world: they reveal the question to be too simplistic. On necessitarianism, perhaps there is exactly one world which is $best_a$ in the strong sense of being better than *all* others, or perhaps there is one or more world which is $unsurpassable_a$ in virtue of being surpassed by no other world. On contingentism, however, there can be no $best_a$ world (simply because there are failures of commensurability between worlds in different clusters), although in principle there could be one or more $unsurpassable_a$ worlds.[14] A world is $best_c$ when it is better than all of its other cluster-mates, and a world is $unsurpassable_c$ when no world in its cluster surpasses it. These terms add nothing new on necessitarianism, since on that view, all worlds are cluster-mates. But on contingentism, perhaps some clusters feature such worlds, while others contain no such worlds.[15] Finally, within a given hierarchy (on either necessitarianism or contingentism) all worlds are both commensurable and comparable, so in principle there could be a world which is $best_h$ by surpassing all of its other hierarchy-mates, or $unsurpassable_h$ in the sense of being unsurpassed by all hierarchy-mates. A complete axiological account will perhaps reveal whether there really are worlds answering to these definitions.

The foregoing can be applied *mutatis mutandis* to universes.[16] Some authors think it obvious that there are unsurpassable universes (McHarry, 1978); others think it obvious that there are not (O'Connor, 2008).[17] But it is important to be careful about what this means.

If necessitarianism about universes is true, then in principle there could be a $best_a$ universe, or $unsurpassable_a$ universes. But if contingentism about universes is true, no universe can be $best_a$, although in principle there could be one or more $unsurpassable_a$ universes. On contingentism, perhaps some clusters features a $best_c$ universe, or one or more $unsurpassable_c$ universes, and perhaps other clusters feature neither kind. Finally, within a given universe-hierarchy (on either necessitarianism or contingentism), all universes are both commensurable and comparable, so in principle there could be a universe which is $best_h$ in the sense of being better than all of its hierarchy-mates, or $unsurpassable_h$ in the sense of being unsurpassed by all hierarchy-mates. Until a more complete axiological of universes is given, it cannot be determined whether there are *best* or *unsurpassable* universes in all these various senses.

The axiological evaluation of universes is particularly important for theistic multiverse theories, since they all posit that there is an objective axiological threshold above which all universes are worthy of inclusion in a multiverse created by God, and at or below which no universe is worthy. Some say little about where this threshold lies. Hudson (2006), for example, asserts that God will create every universe 'worth having' (p. 167): all those which 'satisfy a certain minimal criterion of value' (p. 170). O'Connor simply says that there is an objective threshold, and adds that it may be vague (2008, p. 158). And I have said that God will include 'all and only those universes worth creating and sustaining' (Kraay, 2010a, p. 363). Others offer more detail. Turner (2003) suggests 'a favourable balance of good over evil' (p. 149), and Monton (forthcoming) offers various construals of what this might mean. But such a threshold is too simplistic: a universe might meet this condition, while nevertheless containing some feature that makes it unworthy of inclusion in a divinely furnished multiverse. Perhaps sensing this, some philosophers have proposed additional requirements. Parfit says there should be no injustice (1991, p. 5), and that each individual's life must be worth living (1992, p. 423). Forrest first says that every individual must have a life in which good outweighs evil (1981, p. 53), and later adds two further restrictions: each creature who suffers must at least virtually consent to it, and must receive ample recompense afterwards (1996, pp. 225–7). Draper says that no individual's life may be bad overall, and that God must be a benefactor to all creatures (pp. 319–20). These criteria may well be plausible, but of course they are only partial specifications of a threshold, since they only pertain to universes inhabited by creatures like ourselves. A more complete account of threshold is needed.

Most defenders of the theistic multiverse maintain that God will create *every* universe above the threshold.[18] (Surprisingly, few explicitly add the plausible restriction that God will create *only* these.) O'Connor, however, *denies* that God will create every universe above the threshold (2008, p. 119), and others are silent or neutral on this issue.[19] O'Connor's denial that God will create every universe above the threshold has been persuasively criticized (Almeida, 2010, p. 304 and Monton, forthcoming). At the very least, one who denies that God will create all universes above the threshold bears the burden of undermining or defeating the natural presumption that, all else equal, creating another good thing makes the ensemble better overall.[20]

Finally, while most theistic multiverse accounts in the literature concentrate on God's *creation* of universes, it is important also to incorporate to the traditional idea that God *sustains* universes in existence.[21] In what follows, then, 'TM' will refer to that theistic multiverse which includes all and only those universes worthy of being created and sustained by God. While a *complete* theory will develop this claim in the various ways identified in this section, some important philosophical issues concerning this theory can be discussed prior to these matters being settled. I turn to these now.

2. Problems and Prospects for TM

In this section, I briefly survey some important objections to TM, and some prospects it is thought to have. The first four threaten the coherence of TM.

1. As we have seen, all theistic multiverse accounts presuppose that there is an objective axiological threshold above which universes are worthy of inclusion, and at or below which they are unworthy. But here is an objection familiar from the literature on the Problem of No Best World: perhaps, for any such threshold one might posit, there is a defensible superior threshold, in which case no putative threshold is *good enough* for an unsurpassable being.[22] Theistic multiverse theories must respond to this objection.
2. Tom Talbott and Peter van Inwagen have both complained to me that TM cannot contain *all* universes worthy of being created and sustained, using the following thought experiment. Consider a person *S* who exists in some universe *U* in TM. Assuming trans-world (and trans-universe) identity, that person also exists in universes contained in other worlds, and it is not unreasonable to suppose that at

least one of these is worthy of being created and sustained. But then the theistic multiverse cannot include *all* worthy universes.[23]

3. Gale and Pruss (2003, p. xxvi) anticipate, and Monton (forthcoming) offers, a different argument for the claim that that no multiverse can include *all* universes worthy of creating and sustaining. Rejecting the Principle of Identity of Indiscernibles, Monton says that God could create *duplicates* of worthy universes. Since for any number of duplicates God could create, God could have created more, Monton says, it is impossible for God to create *all* worthy universes.[24]
4. Pruss has suggested that whether a universe is worthy of creation and sustenance may depend upon which *other* universes are worthy.[25] Consider, for example, a pair of universes U1 and U2 which are both very close to the threshold. Suppose that thousands of philosophers and scientists in U1 justifiably believe that U2 does not exist, and that in U2 thousands of philosophers and scientists justifiably believe that U1 does not exist. Suppose, further, that if the relevant belief in U1 is true, the value added thereby to U1 suffices to bring it above the threshold, in which case U2 is below the threshold – and vice versa. On this view, it can seem an arbitrary matter which universe is included.[26]

Apart from those arguments which claim that TM is *incoherent*, three other arguments against TM are worth briefly noting.

5. Appealing to Ockham's razor, someone might hold that multiverses objectionably inflate our ontology. While this charge might be a legitimate complaint against an ad hoc appeal to many independent universes, defenders of the *theistic* multiverse can reply that their model is not ad hoc: it is defended by arguments about what a perfectly good creative agent would do.[27] Moreover, the recent growth in scientific theorizing about multiverses should diminish, if not entirely remove, the initial appeal of this objection.
6. Theists have traditionally held that God's decision to create is a free one, but if one thinks that TM is the unique best of all possible worlds, it might seem that God cannot do otherwise than to select it, and that this in turn compromises God's freedom.[28] And if God's choice is unfree, one might further urge, God is unworthy of thanks and praise for creating.[29]
7. Relatedly, some have held that if TM is the only world that God can choose, it is in fact the only possible world *simpliciter* – a consequence held to be absurd.[30]

These, then, are some of the problems that TM faces. But it is also thought to have some important prospects in the philosophy of religion:

1. As mentioned in the introduction, scientific multiverse theories are thought to undermine the fine-tuning argument. But if philosophers can show that *if theism is true, the actual world is (probably) a multiverse*, this objection will be more difficult to sustain.
2. If TM really is, as some of its proponents claim, the best of all possible worlds, then an important argument for atheism is evaded: the Problem of No Best World.[31]
3. While the problem of *actual* evil holds that evil in the actual world disconfirms theism, the problem of *possible* evil holds that the existence of bad possible worlds disconfirms theism. But if TM is both good and the *only* possible world, then there simply are no bad possible worlds to which the defender of this argument can appeal.[32]

These three prospects have seen relatively little discussion in the literature. In contrast, rather more attention has been paid to a different putative prospect: the claim that theistic multiverse theories can help theists to respond to the problem of evil. In section 3, I outline (and criticize) what defenders of this position have said. In section 4, however, I take issue with some of their critics. In section 5, I argue that multiverse theories do not advantage either side in this debate: they merely require *reframing* certain arguments from evil.

3. Defenders of Multiverse-Based Responses to Arguments from Evil

Of course, there is no such thing as *the* problem of evil. Instead, there are many different arguments for atheism, having different logical forms, which involve one or more premises about evil. Authors who think that a multiverse will help theists typically have one or both of the following argument schemas in mind. (These come in both deductive and inductive variants, and the second premise can be supported in various ways.)

Argument I

1. If theism is true, the actual world is the best of all possible worlds.
2. (Probably,) the actual world is not the best of all possible worlds.

∴ 3. (Probably,) theism is false.

Argument II

1. If theism is true, there is no gratuitous evil.
2. (Probably,) evil *e* is gratuitous.

∴ 3. (Probably,) theism is false.

McHarry asserts that his argument would, if sound, 'dissolve' the problem of evil, which he takes to be Argument I (1978, p. 134). He begins by identifying the best of all possible worlds with a multiverse comprised of *all* (and, we should add, *only*) the threshold-surpassing universes. McHarry grants premise 1, and so wants to deny premise 2. One might expect him to offer evidence for thinking that the actual world *is* such a multiverse,but he does no such thing.[33] More modestly, one might expect McHarry to argue that *our* universe is (probably) above the threshold – or, at least, to block reasons for thinking that it is *not* (probably) above the threshold. McHarry does neither. So McHarry does not 'dissolve' even Argument I, and certainly not 'the' problem of evil.[34]

Turner (2003) advertises his multiverse theory as a 'solution' (p. 143) to the problem of evil, which he takes to be Argument I. But he later moderates this claim, calling it a 'partial answer' and a 'partial solution' (p. 157). Like McHarry, Turner thinks that his multiverse is the best of all possible worlds, and he appears to concede premise 1 – so he presumably wants to block premise 2. Turner says that if God were to create all (and, we should add, *only*) those universes which have a favourable balance of good over evil, 'it should not be surprising if it seems to us that we are in a [universe] that could be better than it is' (p. 157). But this move only blocks *one* reason for thinking that our universe should not be included in the multiverse. Perhaps there are others. Someone might defend premise 2, not by saying that our universe is *surpassable*, but by urging that it contains some feature which God would not permit. Turner continues: '[o]nly if this were a [universe] with a preponderance of evil over good should we conclude that it was a [universe] a benevolent God would not have created' (p. 157). But this is not so. To see why, recall the additional restrictions proposed by Parfit, Forrest and Draper on a universe being worthy of creation and sustenance. If any such condition is plausible, a universe with equal amounts of good and evil (or indeed with *more* good than evil), and in which such a condition is *not* satisfied, would be *unworthy* of inclusion in a theistic multiverse. So, *contra* Turner, there being a preponderance of evil over good is not the *only* way for a universe to be deemed unworthy. At most, then, Turner has

provided a theistic response to just one defence of one premise of one argument from evil.

Hudson (2006) thinks that his *plenitudinous hyperspace* can be deployed against both arguments for atheism displayed above.[35] Hudson thinks that his multiverse is the best possible world, but like McHarry and Turner, offers no reason for thinking that *our* universe is indeed above the threshold he posits: 'worth creating'. So Hudson does not, in this sense, defeat the first argument for atheism. He does, however, suggest a way to use his hyperspace to block a certain defence of premise 2 in a deductive version of Argument II. That defence appeals to considerations about suffering experienced by non-human animals that is *not* caused by agents. Since these considerations can also be used to defend premise 2 of a deductive version of Argument I, Hudson's move is best seen as a response to both.

Hudson takes himself to have an *undercutting defeater* for such a defence of premise 2 of both arguments. In the spirit of St Augustine, he believes it is metaphysically possible that *no* such evil in our universe is gratuitous, since *every* instance is required to bring about some otherwise-unobtainable *aesthetic* good(s) which outweigh(s) the evil in question (2006, pp. 172–81). This response has been criticized in some mistaken ways. Gilmore (2006) and Rea (2008) note that Hudson's move does not require hyperspace: one might instead posit lower-dimensional aesthetic properties as possible justifications for natural evil. This is perhaps true, but of course Hudson did not devise his hyperspace simply to reply to arguments from evil. He takes himself to have independent *a priori* grounds for the claim that, on theism, God would create such a hyperspace. Rea offers another criticism. He points out that all of the allegedly gratuitous evils in 3-space are also contained in 4-space, so 'every reason we have for thinking that our 3-space isn't the best possible will also be a reason for thinking that our 4-space isn't the best possible' (p. 450). But this misses the point of Hudson's argument. Hudson wants to concede that our 3-space may be surpassable, while insisting that this fails to show that the *actual world* is surpassable.

A better criticism is anticipated by Hudson (2006, p. 179), and explicitly levelled by Monton (forthcoming): 'an omnibenevolent being would not value aesthetic properties over preventing an innocent creature from pointlessly suffering'. Monton does not argue for this claim, thinking it obvious. My sympathies are with Monton here: it is difficult to see how a perfectly good being could permit creatures to suffer for the sake of aesthetic goods *which they cannot appreciate*. Now it's true that Hudson merely wants to establish a *metaphysical possibility*, in response

to deductive variants of these arguments. So even if one grants Hudson this much, *and* grants that this is sufficient to undermine this particular defence of premise 2 in both arguments, it should be stressed that this is only a very limited contribution to debate about evil.[36]

4. Critics of Multiverse-Based Responses to Arguments from Evil

So far, we have seen that theistic multiverses should not be thought to 'solve' 'the' problem of evil. In this section, I turn to the critics of multiverse-based responses to arguments from evil.[37] The first two deny that a multiverse containing an infinity of universes above some threshold can be improved by adding another such universe.

Perkins (1980) follows McHarry (1978) in assuming that there are 'optimal' universes as well as 'non-optimal' ones. Perkins assumes that our universe is of the latter sort. But if there are infinitely many non-optimal universes worth creating, he says, 'it is difficult to understand in what sense of "better" the [world] is better for including our [universe]' (p. 170). He doubts that it would be meaningful to say that a world featuring our universe has *more value* than a world without it (pp. 170–1). In his response to O'Connor (2008), Almeida (2010) considers whether God would actualize a super-universe SU_∞ containing infinitely-many universes above threshold τ. Almeida points out that if such a world has infinite value, then removing some or even many universes from it will not diminish its value: 'even if there were infinitely many universes is SU_∞ that included the immense suffering and disvalue found in the actual world, we could remove all of those universes without diminishing the overall infinitely positive value of SU_∞' (p. 306). So, although he appears to grant that our universe has 'on-balance positive value' (p. 306), he thinks that God, on O'Connor's model, would not create it.[38]

Perkins and Almeida thus both believe that God can create an infinitely-valuable universe without including our universe, *and that accordingly*, God should not include it. But there are reasons for resisting this inference, even if the premise is accepted. First, theists sometimes maintain that God is infinitely valuable, in which case the world in which God creates *nothing* might well be thought infinitely valuable. This is the 'bare world': it consists of God, whatever other necessary existents there are, and whatever uncreated contingent beings it contains. Since most theists maintain that at least some worlds in which God creates and sustains contingent entities are better than the bare world,

there is some motivation for thinking that infinitely valuable states of affairs can be bettered.[39]

Almeida himself offers a different reason for resisting this inference (2008, p. 156). Following Vallentyne and Kagan (1997), he imagines a world w_1 which comprises infinitely-many temporal locations, each of which has axiological status 10. Surely, Almeida says, w_1 is on balance better than w_2, a world which comprises infinitely many temporal locations having axiological status 1. Almeida doesn't say what it is for a temporal location to have an axiological status, but this needn't detain us. His point is presumably intended to apply to multiverses as follows: surely it is intuitive to think that multiverse m_1 is *far better* than multiverse m_2 when all the universes in the former are (say) ten times more valuable than all the universes in the latter – even though both multiverses have infinite value.

Both of these reasons motivate searching for a principled way to distinguish the relative value of infinitely good worlds. And, if such a way can be found, this can be used in a response to Perkins (1980). Several such ways are available. First, as Almeida himself notes, 'There are of course nonstandard mathematical representations of infinite value according to which addition and subtraction of infinite numbers is well defined' (2008, p. 156). Here is a second approach. Monton, inspired by Almeida (2008) and Vallantyne and Kagan (1997), suggests the following principle:

> If world w1 has all the locations that w2 has, but w1 has more locations as well, and if the values of all the shared locations are the same, and the values of the non-shared locations in w1 sum to a positive number, then w1 is better than w2.

As Monton rightly observes, a defect of this principle is that it would – implausibly – license God to create some unworthy universes in w_1. But this principle can easily be modified to prevent this unwanted outcome. Consider:

> If multiverse m1 includes infinitely-many threshold-surpassing universes (and no other universes), and multiverse m2 includes all the universes that m1 includes, and also includes threshold-surpassing universes that m1 lacks, (and no other universes) then, ceteris paribus, m2 is better than m1.

This captures the plausible intuition needed to defeat the objection expressed in Perkins (1980) and Almeida (2010).

But suppose that one rejects these moves, or even despairs of finding a principled way to distinguish the relative axiological status of such multiverses. A different approach might concede that (at least some) pairs of infinitely-membered multiverses are *equal in value*, while denying that the value of the two relevant world-actualizing *actions* is equivalent. For example, one might hold that, in such cases, world-actualizing action A1 is better than world-actualizing action A2 just in case the world that results from A1 contains one or more threshold-surpassing universes that the world resulting from action A2 lacks, and no other universes.

While Perkins (1980) and Almeida (2010) deny that a multiverse containing an infinity of universes above some threshold can be improved by adding another such universe, Monton (forthcoming) takes the opposite tack: he holds that God *can* improve such a universe. Monton posits that our universe contains gratuitous evil, but nevertheless is above his axiological threshold: it has a favourable balance of good over evil.[40] One might expect Monton to say that since our universe is above this threshold, it is worthy of inclusion in the multiverse, and that since including it would add to the goodness of reality, then God should include it. Surprisingly, however, this is not what Monton thinks. He says that God could add to the goodness of reality *even more* by creating *duplicates* of *better* threshold-surpassing universes: those lacking gratuitous evil. And so, Monton thinks, God would never 'feel compelled' to create any universe like our own. But this is implausible. If, as he explicitly claims, *every* threshold-surpassing universe adds to the goodness of reality, God has sufficient reason to create *each* one.[41]

5. The Way Forward: Reframing Arguments from Evil

In sections 3 and 4, I objected to claims made by both defenders and critics of multiverse-based responses to the problem of evil. I now argue that multiverse theories are best thought of as requiring defenders of certain arguments from evil to *reframe* their arguments. I begin with an illustrative exchange between O'Connor and Oppy.

O'Connor is more modest than the authors considered in section 3: he merely says that his multiverse 'has some relevance to' the problem of evil (2008, p. 122; 2010a, p. 271). Specifically, he thinks that

> God will in fact have compelling reasons to create a universe in which significant suffering is permitted to occur *even if the goods that require suffering are not the greatest goods*, or *if the universe in which they occur does not belong to a class of supremely valuable realms*. All

> that is required is that the suffering-risking universe satisfy a minimal threshold of goodness. (2008, p. 123; emphasis in original, and see 2010b, p. 316)

It's not clear what the point of O'Connor's first italicized suggestion is: I know of no arguments from evil which claim that God can permit suffering only for the sake of 'the greatest' goods. That aside, O'Connor seems to be saying that those who offer arguments from evil against multiverse-theism bear the burden of showing that our universe fails to surpass the relevant threshold of goodness. Oppy (2008) expresses sympathy for O'Connor's project, but offers the following caution:

> it is perhaps worth pointing out that O'Connor's theory clearly doesn't help to overturn the thought that a morally perfect being would not permit a person to suffer *horrendously* unless it was in the interests of that very person to undergo the horrendous suffering in question. Appearances generated by O'Connor's discussion of the value of universes to the contrary, considerations about minimum thresholds of goodness in suffering-risking universes cannot float free of these kinds of deontological concerns (or so it seems to me).

Oppy is challenging O'Connor to say more about the threshold: in particular, he is asking O'Connor to add a restriction similar in spirit to those offered by Parfit, Forrest and Draper. O'Connor could do so, and then, if he wished to develop his argument further, he could offer reasons for thinking that this restriction is (probably) satisfied in our universe, or at least could try to block reasons for thinking that it is (probably) not.

What this exchange illustrates is that defenders and critics of multiverse-theism can work together to find a mutually agreeable threshold above which universes should be deemed worthy of inclusion in a theistic multiverse. They may well subsequently differ on the *a posteriori* question of whether *our* universe surpasses the agreed upon threshold. Considerations about the existence, variety, magnitude, duration, scope, distribution, types or intensity of evil in our universe could be appealed to in defending the claim that our universe is (probably) *not* worthy of inclusion. And defenders of multiverse-theism could either try to show that our universe (probably) *is* worthy of inclusion, or – more modestly – they could try to defeat or undermine arguments to the contrary. In short, the typical moves in the debate concerning the problem

of evil can easily be *reframed* to apply to multiverse-theism. But this, by itself, will not furnish an advantage to either side.[42]

6. Conclusion

The remarkable recent development of scientific multiverse theories should be taken into account by philosophers. Those who wish to argue that *if theism is true, the actual world (probably) is a multiverse* should take care to ensure that their account is consistent with the best contemporary physics and cosmology. An adequate axiological theory for both worlds and universes is needed, and scientific considerations may constrain this in various ways. Theistic multiverse theories face several important problems, but also offer significant prospects for various issues in the philosophy of religion. Solving the problem of evil, however, is not among these prospects.

Notes

1. For surveys of these and other theories, see: Leslie, 1989; Rees, 2001; Kaku, 2005; Vilenkin, 2006; Carr, 2007; Gribbin, 2009; and Greene, 2011.
2. McHarry, 1978; Parfit, 1991, 1992; Stewart, 1993; Forrest, 1981, 1996; Leslie, 1989; Turner, 2003; Draper, 2004; Hudson, 2006; Collins, 2007; O'Connor, 2008; and Kraay, 2010a and 2011a.
3. Page, 2010.
4. In some theistic multiverse theories, universes are insufficiently distinguished from possible worlds. See, for example, Vallicella's criticisms (1994) of Stewart, 1993, and Almeida's criticisms (2008, pp. 146–8) of Turner, 2003.
5. See Leslie, 1989, pp. 66–9.
6. Interestingly, Lewis concedes in two places that he would prefer not to deny that there can be disconnected space-times (1986, pp. 71, 74). Bricker (2001) seeks to adapt Lewisian modal realism to allow this.
7. Michael Almeida's *theistic modal realism* (2008, 2011).
8. He does define a *universe type* a few pages earlier, but this does not seem to be a *value type* (p. 116).
9. For more on this, see Kraay, 2008.
10. See Kraay, 2010a and 2011b.
11. Gale and Pruss mention this claim (2003, p. xxvii).
12. I discuss this in Kraay, 2011b.
13. In what follows, I assume that every world belongs to a cluster, and that no world can belong to more than one cluster: for every world, there is one unique set of WGMPs and WBMPs which constitutes the appropriate criteria for assessing that world. I make the same assumption concerning universes and their axiological properties.
14. If, on contingentism, there is a world which is is commensurable with no other worlds, it must be deemed (trivially) *unsurpassable*$_a$. The same

consequence applies, *mutatis mutandis*, to the other senses of *unsurpassable* distinguished below.

15. Even if there are failures of *comparability* within a given cluster, this by itself does not preclude there being a $best_c$ world in that cluster. Conversations with Graham Oppy, Ed Wierenga and Yujin Nagasawa helped me to see this.
16. Most theistic multiverse theories appear to assume that all universes are both commensurable and comparable. One exception is O'Connor (2008), who says that universes belonging to different kinds or types are 'incommensurate' (p. 120), or 'likely not fully commensurate' (p. 117). It's not clear, however, whether O'Connor has in mind (using my terminology) failures of *commensurability* or of *comparability*.
17. O'Connor offers a brief argument for this claim (2008, pp. 117–18), but does not defend key aspects of it, including his distinctions between *intensive*, *aggregative* and *organic* values, and his claim that God alone can possess infinite intensive and organic value.
18. McHarry, 1978, p. 134; Parfit, 1991, p. 5; 1992, p. 423; Turner, 2003, p. 149; Hudson, 2006, p. 167; Kraay, 2010a, pp. 361–3.
19. Stewart, 1993; Forrest, 1981, 1996; Draper, 2004.
20. I say more on this in section 4.
21. I discuss this in Kraay, 2010a.
22. I discuss this issue in the context of *worlds* in Kraay, 2010b.
23. It's worth noting that this objection could be expressed in counterpart-theoretic terms. My tentative reply is developed in Kraay, 2011a: there is reason to think that on theism, TM is the *only* possible world, in which case *a fortiori* no individual can inhabit multiple worlds, and no individual has an other-worldly counterpart.
24. Monton's view is puzzling: since many proponents of theistic multiverses already concede that there may be infinitely many universes worthy of being created and sustained, it's not clear what his appeal to *duplicates* adds to the discussion. Perhaps, then, his real worry is that there can be no actual concrete infinities.
25. http://alexanderpruss.blogspot.com/2008/02/complication-for-multiverse-theories_04.html.
26. One might dispute the coherence of the objection in two ways, both of which require defence. First, one might deny that there could be no *other* reason for including (or failing to include) such universes. Second, one might deny that many peoples' justified belief being *true* can raise the overall axiological status of a world in the stipulated way.
27. O'Connor, 2008, p. 122, also makes this point.
28. Almeida (2008, p. 162) levels this charge against Hudson (2006). Those who maintain that God needn't create *every* world above the threshold may preserve some scope for divine freedom here, but at a cost: God's choices on this view have been alleged to be *arbitrary* (Mawson, 2009).
29. I discuss both issues in Kraay, 2008.
30. Almeida (2008) levels this charge against Turner (2003) and Hudson (2006); Monton (forthcoming) responds. I concede the charge in Kraay, 2011a, while tentatively arguing that it may be less serious than it appears.
31. I survey the literature on this argument in Kraay, 2010b.
32. I attempt such an argument, and survey the costs, in Kraay, 2011a.

33. In fact, he says that to demand such evidence is to assume that his proposal is not meaningful unless it can be verified (McHarry, 1978, p. 134). But this is a mistake; one can easily grant that a multiverse model is meaningful, and still ask whether there are good reasons to believe that it corresponds to reality.
34. Similar issues arise for Parfit, who says that that his multiverse model is 'a partial answer' to 'the problem of evil' (1991, p. 5, and see 1992, p. 423).
35. Almeida offers a succinct definition: 'A plenitudinous hyperspace is a collection of many independent three-dimensional subregions in a connected four-dimensional manifold' (2008, p. 158).
36. Megill (2011) offers a unique and interesting meta-argument: he claims that the bare epistemic possibility of there being multiple universes can be shown to defeat *all* arguments from evil. Space does not permit addressing it here; I criticize it in Kraay, forthcoming.
37. Due to space constraints, I do not discuss Draper, 2004.
38. Walker (2009) defends what he calls 'the anthropic argument for atheism', which is similar in structure to certain arguments from evil. He imagines a multiverse-based response to his argument, and replies in the same vein as Perkins and Almeida.
39. Coughlan (1987, pp. 17–18) mentions this move, and Turner (2003, p. 149) appears to endorse it.
40. It is worth noting that many philosophers would deny that a universe featuring gratuitous could be worthy of inclusion in a theistic multiverse, even if it featured a preponderance of good over evil.
41. Of course, as we saw in section 2, Monton thinks it is not possible for God to create *all* threshold-surpassing universes.
42. I am grateful to the Monash University philosophy department for hosting me as a departmental visitor from February–July 2011. Special thanks are due to Graham Oppy for illuminating discussions of key issues, and to Yujin Nagasawa for helpful comments on an earlier draft. I am also grateful to the Templeton Foundation and to Oxford University for supporting my research in 2011–12.

References

Almeida, M. (2008) *The Metaphysics of Perfect Beings* (New York: Routledge).

Almeida, M. (2010) 'O'Connor's Permissive Universe', *Philosophia Christi* 12: 296–307.

Almeida, M. (2011) 'Theistic Modal Realism?', *Oxford Studies in Philosophy of Religion* 3: 1–15.

Bricker, P. (2001) 'Island Universes and the Analysis of Modality', in G. Preyer and F. Siebelt (eds), *Reality and Humean Supervenience: Essays on the Philosophy of David Lewis* (Lanham: Rowman and Littlefield), pp. 27–56.

Carr, B. (ed.) (2007) *Universe or Multiverse?* (Cambridge: Cambridge University Press).

Collins, R. (2007) 'The Multiverse Hypothesis: A Theistic Perspective', in B. Carr (ed.), *Universe or Multiverse?* (Cambridge: Cambridge University Press), pp. 459–80.

Coughlan, M. J. (1987) 'Must God Create Only the Best Possible World?', *Sophia* 26: 15–19.
Deutsch, D. (1997) *The Fabric of Reality* (Harmondsworth: Penguin).
Draper, P. (2004) 'Cosmic Fine-Tuning and Terrestrial Suffering: Parallel Problems for Naturalism and Theism', *American Philosophical Quarterly* 41: 311–21.
Everett, H. (1957) ' "Relative State" Formulations of Quantum Theory', *Reviews of Modern Physics* 29: 454–62.
Forrest, P. (1981) 'The Problem of Evil: Two Neglected Defences', *Sophia* 20: 49–54.
Forrest, P. (1996) *God without the Supernatural* (Ithaca: Cornell University Press).
Gale, R., and A. Pruss (eds) (2003) *The Existence of God* (Aldershot: Ashgate).
Gilmore, C. (2006) 'Review of Hudson, *The Metaphysics of Hyperspace*', *Notre Dame Philosophical Reviews*; available online at http://ndpr.nd.edu/review.cfm?id=7783.
Greene, B. (2011) *The Hidden Reality: Parallel Universes and the Deep Laws of the Cosmos* (New York: Knopf).
Gribbin, J. (2009) *In Search of the Multiverse* (London: Penguin).
Hudson, H. (2006) *The Metaphysics of Hyperspace* (Oxford: Oxford University Press).
Kaku, M. (2005) *Parallel Worlds: A Journey through Creation, Higher Dimensions, and the Future of the Cosmos* (New York: Random House).
Kraay, K. (2008) 'Creation, World-Actualization, and God's Choice among Possible Worlds', *Philosophy Compass* 3: 854–72.
Kraay, K. (2010a) 'Theism, Possible Worlds, and the Multiverse', *Philosophical Studies* 147: 355–68.
Kraay, K. (2010b) 'The Problem of No Best World', in C. Taliaferro, P. Draper and P. Quinn (eds), *A Companion to Philosophy of Religion*, 2nd edn (Oxford: Wiley-Blackwell), pp. 481–91.
Kraay, K. (2011a) 'Theism and Modal Collapse', *American Philosophical Quarterly* 48: 361–72.
Kraay, K. (2011b) 'Incommensurability, Incomparability, and God's Choice of a World', *International Journal for Philosophy of Religion* 69: 91–102.
Kraay, K. (forthcoming) 'Megill's Multiverse Meta-Argument', *International Journal for Philosophy of Religion*.
Leslie, J. (1989) *Universes* (New York: Routledge).
Lewis, D. (1986) *On the Plurality of Worlds* (Oxford: Blackwell).
Linde, A. D. (1986) 'Eternally Existing Self-Reproducing Chaotic Inflationary Universe', *Physics Letters* B175: 395–400.
Mawson, T. (2009), Review of Timothy O'Connor's *Theism and Ultimate Explanation*, *Religious Studies* 45: 237–41.
McHarry, J. D. (1978) 'A Theodicy', *Analysis* 38: 132–4.
Megill, J. (2011) 'Evil and the Many Universes Response', *International Journal for Philosophy of Religion* 70: 127–38.
Monton, B. (forthcoming) 'Against Multiverse Theodicies', *Philo*.
O'Connor, T. (2008) *Theism and Ultimate Explanation: The Necessary Shape of Contingency* (Melbourne: Wiley-Blackwell).
O'Connor, T. (2010a) 'Theism and the Ultimate Explanation: The Necessary Shape of Consistency', *Philosophia Christi* 12: 265–72.
O'Connor, T. (2010b) 'Is God's Necessity Necessary? Replies to Senor, Oppy, McCann, and Almeida', *Philosophia Christi* 12: 308–16.

Oppy, G. (2008) 'Review of O'Connor, *Theism and Ultimate Explanation*', *Notre Dame Philosophical Reviews*; available online at http://ndpr.nd.edu/review.cfm?id=13406.

Page, D. (2010) 'Does God So Love the Multiverse?', in M. Stewart (ed.), *Science and Religion in Dialogue: Volume One* (Oxford: Blackwell), pp. 380–95.

Parfit, D. (1991) 'Why Does the Universe Exist?,' *Harvard Review of Philosophy* (Spring): 4–5.

Parfit, D. (1992) 'The Puzzle of Reality: Why Does The Universe Exist?,' *Times Literary Supplement* (3 July), reprinted in P. van Inwagen and D. Zimmerman (eds), *Metaphysics: The Big Questions* (Oxford: Blackwell, 1998), pp. 418–27.

Perkins, R. K. (1980) 'McHarry's Theodicy: A Reply', *Analysis* 40: 168–71.

Rea, M. (2008) 'Hyperspace and the Best World Problem', *Philosophy and Phenomenological Research* 76: 444–51.

Rees, M. (2001) *Our Cosmic Habitat* (Princeton: Princeton University Press).

Smolin, L. (1997) *The Life of the Cosmos* (Oxford: Oxford University Press).

Steinhardt, P., and N. Turok (2007) *Endless Universe: Beyond the Big Bang* (New York: Doubleday).

Stewart, M. (1993) *The Greater Good Defence: An Essay on the Rationality of Faith* (New York: St Martin's Press).

Tegmark, M. (2007) 'The Multiverse Hierarchy', in B. Carr (ed.), *Universe or Multiverse?* (Cambridge: Cambridge University Press), pp. 99–125.

Turner, D. (2003) 'The Many-Universes Solution to the Problem of Evil', in R. Gale and A. Pruss (eds), *The Existence of God* (Aldershot: Ashgate), pp. 1–17.

Vallentyne, P., and Kagan, S. (1997) 'Infinite Value and Finitely-Additive Value Theory', *The Journal of Philosophy* 94: 5–26.

Vallicella, W. (1994) 'Review of M. Stewart, *The Greater-Good Defence*', *Religious Studies* 30: 243–56.

Vilenkin, A. (2006) *Many Worlds in One: The Search for Other Universes* (New York: Hill and Wang).

Walker, M. (2009) 'The Anthropic Argument against the Existence of God', *Sophia* 48: 351–78.

Part IV
Religious Beliefs

This page intentionally left blank

7

How Relevant Is the Cognitive Science of Religion to Philosophy of Religion?

David Leech and Aku Visala

1. Introduction

On a standard view of the scope of philosophy of religion, philosophers of religion are concerned with the meaning of religious claims and what (if anything) they succeed in saying about the world. Naturalistic explanations of religion, on the other hand, seek to causally explain how religious ideas, beliefs and behaviours come about and persist in human populations.

It is useful to distinguish three broad types of naturalistic explanations of religion:

1. *Evolutionary explanations* that seek to explain the emergence of religion in terms of evolutionary pressures.
2. *Cognitive explanations* that seek to explain why religious ideas are so widespread in human populations in terms of cognitive systems that process them.
3. *Co-evolutionary explanations* that seek to explain the dynamics of religious traditions in terms of cultural evolution.[1]

In this chapter we will be focusing on explanations along the lines of (2) and (3) since researchers identifying with what is now called 'the cognitive science of religion' (CSR) put the stress on cognitive factors in religious belief formation and transmission. However, elements from all three are de facto united (although with differing emphases) in the theory building that goes on under the name of 'CSR'.[2]

Are the results of the CSR relevant for philosophy of religion? As just suggested, we should first make a preliminary distinction between

causes of beliefs and reasons for believing that some propositions are true. Famously, philosophers caution against the danger of *genetic fallacy* in these contexts: we cannot infer from the causes of our beliefs to the truth or falsity of the propositions that are the objects of our belief.[3]

In what follows we will be focusing on the truth-value and rationality of belief in God and how CSR might be relevant for it. We acknowledge that this does not exhaust the scope of the philosophy of religion and that CSR might be relevant for other aspects of philosophy of religion which we will not be discussing here.

2. An Argument for Atheism from the CSR

First let us recall the basic features of the CSR position. The starting point of the approach is that the principal functions of the human mind do not vary across cultures. This is due to the fact that human minds emerge from similar biological foundations (brains) in basically uniform natural environments. Further, our cognitive systems shape or 'constrain' the content of our beliefs and perceptions: our minds select and transform information. Due to such processes, human cognition produces recurrent patterns in human thinking and behaviour across cultures by constraining and informing possible ways of thinking and acting. According to CSR theorists, this is relevant to religion because the most salient religious beliefs (e.g. beliefs about morally interested gods and spirits) can be explained by evoking these systems.

It seems to us that two types of arguments against theistic belief derived from this account have been proposed:

1. Arguments that infer from the existence of a sufficiently complete causal explanation of religious belief (CSR) to the conclusion that there is no God (*no logical space arguments*).
2. Arguments according to which the CSR shows that belief in God is not a product of a belief-forming process that is sufficiently truth-tracking and religious believers are, therefore, not rational in assenting to theistic belief (*unreliability arguments*).

Notice that arguments of type 1 claim that God does not exist whereas arguments of type 2 claim that it is irrational to believe in God.

Although it is often suggested that the CSR will make atheism more probable than theism or even debunk theism in some profound way, the arguments are seldom spelled out in any detail. Consider, for instance, archaeologist Steven Mithen's position. It is a fact, Mithen notes, that

religiosity is a ubiquitous feature of human societies. This fact can be explained in two different ways. (1) There exists a supernatural realm inhabited by a supernatural agent (or agents) that has the power to act in the natural world or even create it. If this explanation is true, then religion is not natural but supernatural and God(s) exist(s). But there is an alternative: (2) the pervasiveness of religiosity may just be due to the fact that the human mind is prone to believe in the supernatural and 'on-going activity of the universe and life are explained by entirely natural processes'.[4] If this were the case, we could safely conclude that 'religion would be simply a curious invention of the human mind'.[5] Mithen concludes that

> religious thought is uniquely associated with *Homo sapiens* and arose as a consequence of cognitive fluidity, which was in turn a consequence of the origin of language. In this regard, there appears to be no need to invoke a moment of divine intervention that initiated the start of a revelation.[6]

This is but one instance of numerous tokens of reasoning that are on the verge of committing the genetic fallacy. The argument seems to be that because we now have a causal explanation for the emergence of religious ideas, we can exclude the possibility of supernaturalism being true in general. The argument is that we now know why and how these puzzling phenomena have evolved and what biological and psychological functions they serve, so we need not think that they represent reality nor assent to their normative claims about how humans should live their lives.

A deductive version of the argument against theism could go something along the following lines.

1. If we have a complete or sufficiently complete causal explanation of how belief in God came about without any reference to the existence of God, then God does not exist.
2. CSR can provide a complete or sufficiently complete causal account of how belief in God came about without any reference to the existence of God.
3. Therefore, God does not exist.

With respect to 2, the claim could be backed up as follows: for the purposes of discovering whether theistic beliefs are justified, it does not matter that we cannot yet explain more recent forms of theistic beliefs

and practices. It is enough if we have an adequate natural explanation of their ancestor forms out of which the more recent forms have grown. This assumes that all current theistic beliefs and practices in the world have a causal ancestry directly involving earlier (indeed, prehistorical) equivalents. At first sight this has a certain plausibility, since until very recently all new religious movements have arisen within societies which were already religious. So the atheist could maintain that it is sufficient for her argument if she can explain naturalistically the prehistoric origin of the basic religious belief disposition.

Let us say that the latter disposition is the disposition to believe that there is at least one supernatural agent, and that we can be in relation with it/them. Adapting a distinction in John Schellenberg, let us call this *basic supernaturalism*, and distinguish it from *elaborated supernaturalism*.[7] The atheist can then argue as follows: we have an explanation of basic supernaturalism, and this is sufficient for establishing that disbelief concerning theism is the most rational response. We do not need to have an explanation of elaborated supernaturalisms, because they are causally dependent on basic supernaturalism. In other words, they would not have arisen if basic supernaturalism had not arisen. And CSR theory provides a sufficient natural explanation of basic supernaturalism.

Now, what reasons do we have to think that 1 and 2 are true? Not very many, we suggest. First of all, it seems that claim 1 commits the genetic fallacy. We simply cannot infer God's nonexistence from the mere fact that there exists a causal explanation for why we believe in God that does not involve God. The causal explanation and God's existence could be compatible; if there is an in-principle incompatibility between the two, this would require further argument. Claim 2 is also quite problematic. The claim is open to what we will call the *Inverted God-of-the-gaps Objection*. We will examine objections to the first premise first and then move on to discuss the second.

3. Objection I: Reasons and Causes

Most philosophers, both atheist and theist, are critical of clear-cut debunking arguments of this sort, and they regard it as a mistake to assume that simply by exposing the causal history of a set of beliefs (even if complete and true) we get to the falsity of the propositions involved. The reason why this way of arguing can be seen as problematic is because we can assess the content of a belief (proposition) independent of the causes which made people hold the proposition true (belief). Every belief has a causal history of some kind and in this sense can to

some extent be explained causally, so the possibility of causal explanation itself cannot count against the truth of a belief. The question is not whether a belief, for instance, that God exists, has causes, but whether those causes are of an appropriate kind. What kinds of causes are allowed or not is a matter of considerable dispute in epistemology. We will return to this point later. Suffice it to say that this dispute is ultimately about the rationality of believing a proposition, not about the truth or falsity of a proposition.[8]

For such reasons, most atheist and theist philosophers do not share Mithen's faith in the power of causal explanations to settle philosophical disputes about the truth or falsity of religion.

4. Objection II: Inverted God-of-the-Gaps Objection

The next objection – the *Inverted God-of-the-gaps objection* (IGO) – has to do with premise 2. In short, IGO consists of two claims. According to the first, it is not true that CSR can provide us with a complete or sufficiently complete causal explanation for belief in God. In other words, the claim is, simply, that we have good reasons to think that the CSR causal account of religious belief is not one which would allow us to exclude the possibility of God being directly involved in the process. The second claim included in IGO is that CSR only explains the emergence of religion in a very generic way rather than explaining any particular individual's belief in God.

First, the theist can claim that the CSR falls short of a complete, or sufficiently complete, explanation. Let us consider the most trivial point first. We simply cannot inspect the original Pleistocene context in which religious behaviour supposedly evolved, and the evidence we have by which we can reconstruct something about it is very slim. We are reduced to informed speculation about what actually happened at the point at which our primate ancestors gradually became us and developed religiosity. So we need to have some way of knowing whether the CSR account is the *best* speculation about how religious belief and practice arose and became stabilized.

The theist defender of IGO can here argue that the mechanisms specified by CSR are not *sufficient* to have produced (and perhaps even maintained) religious beliefs and practices in the Pleistocene, or at least that we do not *know* whether they are sufficient. She can point out that CSR fails to explain, for instance, why *specific* individuals have the kind of beliefs they have about God. She can also point out that CSR does not account for religious experiences and the details of more

sophisticated forms of religion. It can, therefore, be claimed that CSR explanation concentrates on general tendencies in religious thinking and behaviour and thus has scope-related, methodological limitations. This generality is mainly a result of the CSR's commitment to a cultural epidemiological approach to its subject matter inspired by Dan Sperber's anthropological work. Sperber's *epidemiology of representations* model, endorsed by CSR researchers, explains population-level tendencies to religiosity rather than individual cases of religious beliefs or behaviours.[9] CSR consequently explains the 'cultural fitness' of representations or their likelihood to spread, accounting for the memorable nature of religious representations by appeal to their fit with the properties of the postulated nonconscious mental tools.[10]

It follows that the CSR explanatory model underdetermines how humans acquire and transmit religious beliefs and behaviours.[11] For example, it is beyond the scope of CSR explanation to account for why Trinitarian belief arose in communities of Christian theists, even though it can predict generally that these communities would believe in a supernatural agent or agents. Further, CSR cannot explain why given individuals hold particular theistic beliefs, and whether those beliefs are rationally justified. Particular theistic beliefs (along with all other religious beliefs) are presumably supported by the cross-cultural human intuitive biases and predilections specified by the CSR. However, they do not all by themselves produce particular theistic beliefs. In other words, it would be misconceived to suppose that CSR explanations specify all of the causal pathways through which people come to hold theistic beliefs.

5. A Probabilistic Version of the Argument

We have now considered some reasons to think that premises 1 and 2 in the previous deductive argument are false. But perhaps if we reformulate the atheistic argument in a probabilistic way, we could circumvent at least some of the objections.

Consider the following version of the argument.

4. According to the CSR, the causes of our belief in the existence of God are mostly natural instead of supernatural.
5. If God exists, it is likely that the causes of our belief in his existence are mostly supernatural.
6. Therefore, God does not exist.

The deductive version of the argument examined in the previous section attempted to show that God does not exist. The probabilistic version only attempts to show that CSR theories give us a reason to prefer atheism over theism.

The theist could respond in various ways here. First, she could grant that the argument successfully increases the probability of atheism over theism. The theist could, nevertheless, say that she has good independent reasons to believe in the existence of God and claim 6 is not enough to lower the probability of theism to such an extent that she should cease believing it. The theist could concede that the causal pathways through which people come to believe in the existence of God are largely natural and this would be something that seems unlikely on the theistic hypothesis. However, she could still invoke other arguments or considerations for the existence of God and maintain that theism does significant explanatory work in other domains (by explaining the existence, rationality and life-permitting qualities of the universe, for instance).

The argument could also be challenged by using the IGO points against premise 4. The opponent of the argument can claim that the CSR in not complete enough to catch the notions of, for example, revelation or personal religious experience. Nor is there anything in the CSR that would give us a reason to rule these possibilities out as causes of a particular individual's belief in God or as a basis of institutional religious practice. The wielder of IGO would therefore conclude that the CSR does not, by itself, give us reasons to conclude that belief in God is 'mostly' (whatever that means) naturally caused.

Premise 5 is a premise that is seldom stated; instead it is usually simply assumed. What is presupposed here is that if God exists, he would probably make his existence known to us and enter into relationship with us in a way which was mostly supernatural, or at least supernatural in some relevant respect that can be observed either scientifically or in some other commonsensical way. No one, not even the most hardened supernaturalist, claims that there are no natural causes at work in causing belief in God. What the defender of 5 would then mean is that there are more natural causes at work than would be expected given theism. That is to say, given theism we would expect God to interact with humans and cause belief in him in a supernatural way. But what reasons do we have to suppose this?

The supposition could be challenged by arguing that God could, as far as we know, use natural causes (e.g. our ordinary cognitive systems, sociological and cultural mechanisms) to supplement supernatural

influence (e.g. action of the Holy Spirit, direct revelation etc.) to a large extent. Since God is, by definition, omnipotent, He would have the power to do this. The question is whether acting this way would somehow be against God's purposes or be in contradiction with his goodness (or some other relevant attribute). There might be a problem for the theist here and we will be returning to it later.

6. Arguments from Unreliability

In addition to the aforementioned arguments, there is another kind of argument that instead of attacking the *truth* of theism, is directed against the *rationality* of belief in theism. Earlier we dubbed these *unreliability argument*s, since they invoke the unreliability of the mechanisms that produce belief in God. Take the following claim by psychologist Paul Bloom as an example:

> No finding from the cognitive science of religion can refute either the existence of God or a theistic account of the origins of religious belief. But, even so, psychological inquiry can still tell us something about the rationality, or lack thereof, of religious believers, in the same sense that it can tell us about the mental status of those who believe in life on other planets.[12]

Unreliability arguments seem at least prima facie plausible: if we can demonstrate that belief in God is completely naturally caused, that is, not in any direct way connected to God, then it seems that our belief that God exists is not formed in a properly truth-tracking way.

Structurally speaking, unreliability arguments follow a common pattern. In the literature, arguments of this kind have been dubbed *evolutionary debunking arguments*.[13] Most arguments of this sort have to do with morality, but the same strategy is now also being employed against belief in God. According to philosopher Guy Kahane, evolutionary debunking arguments are of the following form:[14]

> S's belief that p is explained by X
> X is a process that does not track the truth
> Therefore, S's belief that p is not justified

From the point of view of debunking arguments, it does not really matter whether God actually exists or not, because the truth of theism is not a necessary condition for the rationality of belief in God.

We can now present a somewhat simplified unreliability argument from the CSR for the irrationality of theistic belief.

7. S's belief that there is a God is explained by CSR (that is, by S's intuitive cognitive mechanisms).
8. The cognitive mechanisms that cause belief in the existence of God are not truth-tracking with respect to the existence or non-existence of God.
9. Therefore, S's belief that there is a God is not justified.

The defender of such an argument can invoke several considerations in support of premise 8. The basic claim would be that in some sense the CSR shows that the cognitive mechanisms responsible for causing our belief in God are not a sufficiently reliable source of beliefs. Here are just a few possible ways to defend this claim.[15]

First, one could claim that there needs to be a proper kind of causal connection between a belief and its target for a belief to be justified. But such a connection between our belief in God and gods seems to be missing, since given CSR there is no need to think that God directly caused humans to believe in him. Second, one could claim that our hypersensitive agency detecting device (HADD) produces false beliefs to such an extent that it should not be regarded as a reliable belief-forming mechanism. If HADD is unreliable and it is indeed central in the CSR account of the emergence of belief in God, then belief in God is not justified. Third, it could be claimed that the mechanisms that CSR identifies produce all sorts of wildly diverse beliefs about God and gods. From this actual diversity of beliefs we can then infer that we should not take such systems as reliable sources of beliefs about God.

Basically, theists have adopted two different strategies in answering such claims. The first consists of demonstrating that there are no reasons to suppose that the mechanisms described by the CSR are unreliable in all religious contexts without begging the question against the theist. With respect to the claim that there is no sufficient causal connection between belief in God and its target, some have argued that CSR does not really show this. For instance, Michael Murray claims that as far as we know, God might be active further back in the causal nexus and set up the initial physical conditions of the universe and guided evolution in such a way that we would develop capacities to form beliefs about him. According to this 'cognitive fine-tuning' view, there is no need for God to intervene by special divine action at every step of religious belief formation and thus the completeness of the CSR account is preserved.[16]

The second argument from the unreliability of HADD can also be challenged. The claim would then be that since HADD is obviously reliable in some contexts and not others, it cannot be shown to be unreliable in religious contexts without first assuming that there is no God. Most of the time our agency detection is correct in attributing agency. Simply on the basis of the CSR, as far as we know, it could also be reliable in the case of God. In addition, it could be insisted that since even CSR writers themselves disagree as to what the explanatory scope and function of HADD actually is, its role in explaining individual instances of belief in God might actually be quite small.[17]

Third, Murray has also argued that the diversity of religious beliefs does not really count against the reliability of the cognitive mechanisms that produce them. This is because the diversity is actually produced by the cultural information that elaborates the basic intuitions produced by the cognitive mechanisms. The mechanisms themselves only produce certain intuitions to which different kinds of elaborated cultural constructions can then attach themselves. Thus, the diversity of religious beliefs does not give us enough reason to think that the underlying systems are unreliable sources of belief in God.

Finally, one could also add that our judgements about reliability or unreliability of belief-forming mechanisms tend to be relatively vague and depend heavily on background assumptions. All this suggests that at least premise 8 of the standard unreliability argument could be challenged.

But there might be another way to respond to arguments based on unreliability. This strategy utilizes the IGO objection that was introduced earlier. The previous objections were based on refuting premise 8 of the standard unreliability argument. The wielder of IGO will instead attempt to refute premise 7 and with it the claim that CSR in fact explains individual belief in God in a sufficient way. Since we have already discussed IGO earlier, we will keep these remarks brief.

If IGO is correct, then it is the case with respect to any given individual's belief in God that CSR underdetermines the causes of that individual's belief in God. As we pointed out earlier, the rationality of theism depends at least to some extent on reasons, arguments and experiences that are not explained by CSR-type mechanisms. These causes cannot be accounted for in terms of non-conscious cognitive systems. If this is the case, then even if CSR turns out to be true and we could demonstrate that the cognitive mechanisms that CSR reveals are unreliable, the theist might still be justified in believing in God on the basis of these other reasons.[18]

We thus conclude that current unreliability arguments along the lines of premises 7–9 have been unable to show that the cognitive systems which more or less cause our belief in God (according to CSR) are so radically unreliable that theistic belief is irrational.

7. The Rationality of Theism and *Sensus Divinitatis*

Attempts to support theism with findings of the CSR have been rare but not inexistent. These attempts have mainly focused on the rationality of theism, rather than its truth. For instance, such an argument is more or less explicit in an article entitled 'What Theology Might Learn (and Not Learn) from Evolutionary Psychology: A Postliberal Theologian in Conversation with Pascal Boyer' by theologian Niels Henrik Gregersen. Gregersen sees no particular theological problem in the claim that religious beliefs are caused by our ordinary, biologically produced cognitive mechanisms. However, he insists (correctly, in our view) that the truth of religious beliefs should not be evaluated in terms of their biological origin.[19]

Gregersen is mainly interested in the origins and function of rationality in different domains of life. If religious rationality has cognitive and evolutionary roots in the sense that it is not its own distinct domain of rationality, all the better, he claims, for belief in God.[20] First, the fact that religious concepts are easy to understand and process coincides with the Judeo–Christian idea that humans are created in the image of God and thus endowed with abilities to be in communication with him. This ability to understand God is common to all humans irrespective of their religious views. So the fact that CSR finds natural mechanisms that explain the pan-human tendency to believe in godlike agents is something that theism would give us reason to expect.

Simply put, Gregersen's argument is that if CSR is correct and religion is indeed a by-product of non-religious cognitive systems (such as HADD and theory of mind), then it is the case that belief in God is generated by the same cognitive mechanisms as most of our non-religious beliefs. If, then, to avoid global scepticism, we should hold the deliverances of our ordinary belief-forming mechanisms prima facie justified, and if belief in God is a product of these ordinary mechanisms, then it is prima facie justified in the same way as all other basic beliefs are.

The problem with proposals such as Gregersen's is that they do not get the theist very far: prima facie justification can be easily removed by defeaters of different kinds (reasons, arguments, new experiences). At most, it could show that religious reasoning *in general* is justified.

But it would do nothing to increase the likelihood of theism being the case.

Psychologist Justin Barrett has made arguments along similar lines. If belief in God is largely produced by normal human cognitive systems and if we generally trust these systems, then we should not suspect them in the case of belief in God. Barrett compares belief in God and belief in other minds and argues that they are psychologically speaking on the same level.[21] Both types of beliefs – belief in gods and belief in other minds – (1) lack empirical support, (2) are believed both reflectively and intuitively and (3) are mostly universal. Barrett points out that it is not scientific evidence that makes people believe in minds but rather the ordinary outputs of their cognitive systems. Thus, the belief in minds and belief in God should be held as foundational or basic. Foundational beliefs – belief produced by our ordinary belief-forming mechanisms – need no outside justification to be prima facie rational. What is needed is that there is no immediate defeater to them. Further, as they are both basic beliefs, if we doubt the one on the basis of lack of propositional evidence we should also doubt the other.

This line of argument and the idea of the prima facie justification of basic beliefs is indebted to Reformed Epistemology. Cognitive theories, according to Barrett, may have ended up revealing a form of what traditional Calvinist theology referred to as the *sensus divinitatis*, or 'God-sense' (a God-given capacity to know God).[22] According to this traditional theological view, the *sensus divinitatis* is the source of our basic, non-inferential beliefs about the nature and existence of God and it is providentially guided by God. *Sensus divinitatis* is something that all human beings have, although some have refused to listen to it.[23]

Notice that those who link *sensus divinitatis* with the CSR-type mechanisms have two different arguments going on at the same time. The first argument, as we have seen, is for the rationality of theism and is based on the Reformed Epistemology account of basic beliefs and their justification. But in addition to this, there is also an argument of a different type – an argument that is rarely spelled out. It would go something like this. The reasons for postulating the existence of *sensus divinitatis* are mainly theological and have been around for a considerable time. If it turns out to be the case that CSR confirms the existence of such a mechanism, then CSR would provide a reason to think that theism is true. In such a case, CSR would function as an independent source of knowledge that would confirm what is being claimed on a theological basis, thus giving support to theism.

Although it might seem promising to link CSR results with the *sensus divinitatis* doctrine, on closer inspection several problems arise. First, there is the problem of the IGO. That is to say, belief in God might be too content-specific to be a product of the mechanisms described by the CSR (recall Murray's defence). The theist would need a lot more than just *sensus divinitatis* to get to full-fledged faith in God. Of course, the theist can claim that *sensus divinitatis* does not give us a fully fledged concept of God in the first place, but rather a vague feeling that 'there might be something out there', to which she can add other sources for beliefs about God.

Second, the CSR only makes it probable that some kind of supernatural belief arises, not inevitable. Some theists might not like the idea of God as a risk-taker. It is true that our cognitive architecture seems to have a survival advantage over some other types of architectures so it is probable that something of the sort would come about through natural selection, but this is by no means certain or necessary. All products of evolution are to some extent contingent. Therefore, if CSR-type mechanisms were identical with *sensus divinitatis*, there is always the chance that *sensus divinitatis* might not be realized, at least not by natural selection.

Third, if God chooses to be in contact with human beings only through the mechanisms described by the CSR, then it would seem that God would not be in communion with humans in any direct way. Most theists, however, believe that God is somehow more intimately related to us, indeed, that they enjoy a personal relationship with him. If we claim that God interacts with us only through the mechanisms that CSR describes, then God would come out looking like a distant and deceiving God.[24] Given these worries, the theist should probably be hesitant in identifying *sensus divinitatis* with CSR-type mechanisms.

8. Objection: The Deceiving God

The last point about the possibility of a deceiving or distant God can be expanded into a more general argument against theism from CSR. We call this the Deus Deceptor (DD) argument. For instance – and this picks up on the intuition behind the probabilistic argument in the previous section – the CSR account might be interpreted as entailing that humans do not relate directly to a supernatural God, but rather to a more or less adequate simulacrum of him produced by their cognitive architectures. But the atheist can object at this point that this causal

account of religious belief is not compatible with God's putative moral nature. If God is supposed to be good, non-deceiving and so forth, then he should not produce belief in us in the way specified by CSR. So if a theist suggests that he does produce human beliefs about him in this way, this should strike the theist and atheist alike as improbable given what we take to know about the nature of God. In that case, the CSR account does not render theism more but *less* probable.

Perhaps the theist could reply at this point that God intervened in human history at a certain point and henceforth the beliefs he put into certain humans' minds were correct ones, and indeed the result of direct divine action. But the atheist can insist that even if this were the case, it would still be incompatible with God's benevolent nature to have allowed vast amounts of false, merely natural, religious beliefs. First, he would have deprived millennia of humans from enjoying meaningful direct loving relationship with him. Second, it would presumably surprise theistic expectations to learn that God at once allowed millennia of humans – who, presumably, were *non*-resistant to the Divine call – to have false merely natural 'religious' beliefs, and selectively put involuntary true religious beliefs into the minds only of some.[25]

The theist can perhaps further object that God may have been giving earlier humans a foretaste of himself through these beliefs which arose merely through the mechanisms specified by CSR. However, she must still face the question of why God did this through simulacra, and did not reveal himself personally even if in a limited way.

Indeed, the non-theist could produce an argument which would confirm prior non-theist assumptions, and thus represent a reversal of the theist's *sensus divinitatis* argument. This argument would go as follows: the non-theist does not find the CSR explanation of religious belief and behaviour improbable in relation to her background assumptions (unguided evolution, comprehensive natural causation, natural human cognitive architectures etc.). By contrast, the traditional theist would (or should) find the CSR explanation of religious belief and behaviour improbable in relation to their background assumption that there is an omnibenevolent, non-deceiving God who would choose to enter into loving, meaningful, direct relationship with humans.

We can see the extent of the problem for the theist more clearly if we see it in terms of the problem of divine hiddenness. According to, for instance, Schellenberg, any evidence for the absence of loving personal relationship between God and humans counts centrally against theism. If theism is the case, a relationship with God will be the ultimate good for humans. If we grant this, then it would look like God's

moral attributes would count as reasons for thinking that God would *not* realize certain states of affairs if he existed.[26]

But the CSR account, as we have seen, can plausibly be interpreted as implying that God, if he exists, allowed very inadequate *simulacra* of relationship with him to exist in human minds *in the place of direct personal relationship* for most of human history. Two possibilities concerning how God might have done this come to mind, neither of which help the case for theism.

First, perhaps God created humans *incapable* of entering into direct conscious relationship with God by creating them without the requisite cognitive and affective capacities to do so. But it is hard to see why a God with the relevant moral properties would have done this, since it renders (non-resistant) humans unable to enter into genuine conscious relationship with God, which is the opposite of what the theist would expect.

Second, perhaps God created humans with the capacity to enter into direct conscious relationship with him, but decided for some reason to replace this with *simulated* relationships for most of human history (including our long prehistory), only enabling a true, deeply meaningful, direct relationship with the advent of Judaeo–Christianity. But then the question suggests itself: if God could have enabled true, deeply meaningful, direct relationship for all humans, why did he not do so? Perhaps here the theist would have to develop some defence of the necessity of humans coming to more and more adequate knowledge of God gradually. Possibly a version of theistic evolutionism applied to humans' cognitive and affective capacities might work here. But the details of this would have to be fleshed out, and in any case one can hardly talk about gradualism here, since the difference between having a simulated experience of God and having a real, direct experience of him is not one of degree.

We might finally mention a further way in which CSR could frustrate theistic expectations which was briefly mentioned above. CSR theory would seem to imply that large-scale human non-theism (or deeply imperfect theism) over countless generations cannot plausibly be blamed on human voluntary resistance to God. The problem can be stated like this: for most of human history and prehistory, humans involuntarily held religious beliefs which were either not theistic ones or very imperfectly 'theistic'. The involuntariness is the problem here for the theist who wants to reconcile her theism with CSR. On the CSR account, these humans cannot plausibly be interpreted as *resisting* deep, meaningful relationship with God and the theistic beliefs which would arise involuntarily from such a relationship. As a matter

of fact – and through no fault of their own, but as a result of their cognitive architectures – they did not possess the appropriate beliefs about a super-person distinct from nature who has the omni-attributes. So not all nonbelief in the theistic God is the result of voluntary resistance to God (i.e. sin).

The basic point here is perhaps this: even if the God of love might accept with regret that humans might resist awareness of him, he would not – if he was perfect – himself make direct, conscious relationship with himself impossible.[27] But the theist who accepts CSR seems to be constrained to say that this appears to be what God has done (at least for prehistory and prior to the Abrahamic revelations).

In conclusion, then, it could be said that there are prima facie good reasons for thinking that while the existence of just any naturalistic causal account of religious belief formation does not make theism more or less likely, certain features of the CSR account look like they could make theism more improbable than not, even while they do not lend active support to atheism.

9. Conclusions

We think the implications of CSR for the philosophy of religion, contrary to how they have often been presented, are very inconclusive. CSR theory can be something of a double-edged sword for atheists and theists hoping to recruit it to support their respective positions. The same CSR-based arguments which seem to favour one side in the debate can, appropriately transformed, favour the other side.

For instance, IGO can be used to show that there are present gaps in naturalistic explanations of religion, but this can also be re-converted by the atheist into a standard God-of-the-gaps objection to theism: the gaps might be explained by the fact that researchers have so far given insufficient attention to this field (this can gain extra plausibility from the fact that the CSR field is new). Likewise, appeals to CSR as confirmatory of *sensus divinitatis* can also be turned against the theist as actually offering reasons for thinking that the new naturalistic explanations of religion might actually disconfirm theological expectations.

So, on the basis of this survey of arguments it seems to us reasonable to say that the CSR fails to offer unequivocal support either to atheism or to theism, and indeed its implications for the atheism–theism debate are somewhat unclear. We therefore suggest that the CSR as it currently stands provides very little or no reason to prefer atheism over theism, but also no reason to prefer theism over atheism.

Notes

1. For an overview of explanations along these lines, see Schloss and Murray, 2009.
2. See, e.g., Lawson and McCauley, 2002; Boyer, 1994, 2001; Barrett, 2004; Whitehouse, 2004; Atran, 2002.
3. By contrast, some prominent authors have recently tried to imply that because an empirically supported naturalistic explanation of religion now exists, religious truth-claims are once and for all debunked. See for instance Dawkins, 2006 and Dennett, 2006.
4. Mithen, 2009, p. 11.
5. Similar suggestions have been made by, e.g., Bering, 2010.
6. Mithen, 2009, p. 27.
7. Schellenberg, 2005, pp. 37–8.
8. Cf. for instance atheist philosopher John Mackie: 'Even an adequate, unified, natural history which incorporated all these factors would not in itself amount to disproof of theism. As William James and many others have insisted, no account of the origin of a belief can settle the question whether that belief is true or not' (1982, p. 197).
9. See, e.g., Sperber, 1996.
10. Indeed, critics have questioned CSR's explanatory power precisely on account of this generality, and suggested that it may be the type of large-scale anthropological programme which has little to give to the study of particular cultural traditions because the generalizations it produces are not very interesting. See Laidlaw, 2007 and Day, 2007.
11. This point is developed further in Leech and Visala, 2011. For a more comprehensive treatment, see Visala, 2011.
12. Bloom, 2009, p. 126.
13. For evolutionary debunking arguments of religion, see Barrett, Leech and Visala, 2010 and Leech and Visala, 2011. See also essays in Schloss and Murray, 2009.
14. Kahane, 2011, p. 103.
15. For arguments and responses see Barrett, 2007; Murray, 2007, 2009. For an overview, see Leech and Visala, 2011.
16. Murray and Goldberg, 2009.
17. On HADD's explanatory scope, see Barrett, 2004, p. 39.
18. This response to unreliability arguments is developed in Leech and Visala, 2011, pp. 180–3.
19. Gregersen, 2006, p. 320.
20. Gregersen, 2006, pp. 312–14.
21. Barrett, 2004, pp. 95–105. Barrett's argument draws from Alvin Plantinga (1990).
22. See, Barrett, 2009, p. 98. The *sensus divinitatis* argument is presented in more detail by Clark and Barrett, 2010.
23. See Plantinga's treatment of *sensus divinitatis* in Plantinga, 2000, ch. 6.
24. These problems are outlined in Leech and Visala, 2011.
25. On the hard problem of non-resistant disbelief for theism, see Schellenberg, 2007, pp. 198–206.

26. Schellenberg, 2007, p. 198.
27. Schellenberg, 2007, p. 202.

References

Atran, S. (2002) *In Gods We Trust: The Evolutionary Landscape of Religion* (New York: Oxford University Press).

Barrett, J. (2004) *Why Would Anyone Believe in God?* (Walnut Creek: AltaMira Press).

Barrett, J. (2007) 'Is the Spell Really Broken? Bio-Psychological Explanations of Religion and Theistic Belief', *Theology and Science* 5: 57–72.

Barrett, J. (2009) 'Cognitive Science, Religion and Theology', in M. Murray and J. Schloss (eds), *The Believing Primate: Scientific, Philosophical, and Theological Reflections on the Origin of Religion* (New York: Oxford University Press), pp. 76–99.

Barrett, J., D. Leech and A. Visala (2010) 'Can Religious Belief Be Explained Away? Reasons and Causes of Religious Belief', in Ulrich Frey (ed.), *The Nature of God – Evolution and Religion*, vol. 1 (Marburg: Tectum Verlag), pp. 75–92.

Bering, J. (2010) *The God Instinct: The Psychology of Souls, Destiny and the Meaning of Life* (London: Nicholas Brealey).

Bloom, P. (2009) 'Religious Belief as an Evolutionary Accident', in M. Murray and J. Schloss (eds), *The Believing Primate: Scientific, Philosophical, and Theological Reflections on the Origin of Religion* (New York: Oxford University Press).

Boyer, P. (1994) *The Naturalness of Religious Ideas: A Cognitive Theory of Religion* (Berkeley: University of California Press).

Boyer, P. (2001) *Religion Explained: The Evolutionary Origins of Religious Thought* (New York: Basic Books).

Clark, K. J., and J. Barrett (2010) 'Reformed Epistemology and the Cognitive Science of Religion', *Faith and Philosophy* 27: 174–89.

Day, M. (2007) 'Let's Be Realistic: Evolutionary Complexity, Epistemic Probabilism, and the Cognitive Science of Religion', *Harvard Theological Review* 100: 47–64.

Dawkins, R. (2006) *The God Delusion* (London: Bantam Press).

Dennett, D. (2006) *Breaking the Spell: Religion as a Natural Phenomenon* (New York: Viking).

Gregersen, N. H. (2006) 'What Theology Might Learn (and Not Learn) from Evolutionary Psychology: A Postliberal Theologian in Conversation with Pascal Boyer', in F. LeRon Shults (ed.), *Evolution of Rationality: Interdisciplinary Essays in Honor of J. Wentzel Van Huyssteen* (Grand Rapids: William B. Eerdmans), pp. 306–26.

Kahane, G. (2011) 'Evolutionary Debunking Arguments', *Nous* 45: 103–25.

Laidlaw, J. (2007) 'Well-Disposed Social Anthropologist's Problem with the "Cognitive Science of Religion" ', in H. Whitehouse and J. Laidlaw (eds), *Religion, Anthropology and Cognitive Science* (Durham: Carolina Academic Press), pp. 211–46.

Lawson, E. T., and R. McCauley (2002) *Bringing Ritual to Mind: Psychological Foundations of Cultural Forms* (Cambridge: Cambridge University Press).

Leech, D., and A. Visala (2011) 'The Cognitive Science of Religion: A Modified Theist Response', *Religious Studies* 47: 301–16.
Mackie, J. L. (1982) *The Miracle of Theism: Arguments for and against the Existence of God* (Oxford: Clarendon Press).
Mithen, S. (2009) 'The Prehistory of the Religious Mind', in N. Spurway (ed.), *Theology, Evolution and the Mind* (Newcastle: CSP), pp. 10–30.
Murray, M. (2007) 'Four Arguments That the Cognitive Psychology of Religion Undermines the Justification of Religious Belief', in J. Bulbulia et al. (eds) *The Evolution of Religion: Studies, Theories, and Critiques* (Santa Margarita: Collins Foundation Press), pp. 365–70.
Murray, M. (2009) 'Scientific Explanations of Religion and the Justification of Religious Belief', in M. Murray and J. Schloss (eds), *The Believing Primate: Scientific, Philosophical, and Theological Reflections on the Origin of Religion* (New York: Oxford University Press), pp. 168–78.
Murray, M., and A. Goldberg (2009) 'Evolutionary Accounts of Religion: Explaining and Explaining Away', in M. Murray and J. Schloss (eds), *The Believing Primate: Scientific, Philosophical, and Theological Reflections on the Origin of Religion* (New York: Oxford University Press), pp. 179–99.
Plantinga, A. (1990) *God and Other Minds: A Study of the Rational Justification of Belief in God* (Ithaca: Cornell University Press).
Plantinga, A. (2000) *Warranted Christian Belief* (New York: Oxford University Press).
Schellenberg, J. (2005) *Prolegomena to a Philosophy of Religion* (Ithaca and London: Cornell University Press).
Schellenberg, J. (2007) *The Wisdom to Doubt* (Ithaca and London: Cornell University Press).
Schloss, J., and M. Murray (eds) (2009) *The Believing Primate: Scientific, Philosophical, and Theological Reflections on the Origin of Religion* (New York: Oxford University Press).
Sperber, D. (1996) *Explaining Culture: A Naturalistic Approach* (Oxford: Blackwell).
Visala, A. (2011) *Theism, Naturalism and the Cognitive Study of Religion: Religion Explained?* (Aldershot: Ashgate).
Whitehouse, H. (2004) *Modes of Religiosity: A Cognitive Theory of Religious Transmission* (Walnut Creek: AltaMira Press).

8

The Rationality of Classical Theism and Its Demographics[1]

T. J. Mawson

I

Whatever the naturalness of the religious impulse, it is obvious that nurture directs it and may either enhance or diminish it. Historians and pre-historians have long argued that varieties of polytheism (and/or ancestor- and spirit-worship) were extremely widespread, if not universal, in the earliest recoverable periods of human development. And it has long been the consensus that monotheism first emerged, fitfully but severally, in what is sometimes called the Axial Age, before gaining eventual hegemony over other supernaturalist views. Finally, there would be little dispute that, with the modern age, naturalism – the view that there is no supernatural order whatsoever – has emerged from the shadows as the view of a significant minority; in some countries it has even managed to become the majority view.

The empirical facts of the demographics of theism, both over time – considering humanity as a whole, as it has developed – and over space – considering the set of humans extant at any one time – have been supposed by some to have implications for the rationality of classical theism. It is these supposed facts and their supposed implications that I wish to explore in this chapter. Obviously facts of the past are harder to discern than facts of the present, *ceteris paribus*. For this reason, in what follows I shall focus on the contemporary demographics of theism; the same points that may be made about these facts may be made *mutatis mutandis* about the facts as we may think we discern them concerning the historical and pre-historical demographics of theism.

In a moment, then, I shall sketch the contemporary demographics of theism as I perceive them, marking the divide between those who subscribe to some variant of theism and those who do not (lumping

together into the latter category both agnostics, understood as those who fail to have a belief either way, and atheists, understood as those who believe that there's not a God). Before I sketch this pattern, it is as well for me to make mention of some worries that one may have with the method that leads me to perceive it, namely – in a manner comparable to much contemporary so-called 'experimental philosophy' – that of aggregating the results of various surveys, polls and responses to questionnaires, to make mention of these worries and to do something to defuse them.[2]

II

It must be conceded that it is certain that people responding to a question asking them whether or not they believe in God understand 'God' in a variety of ways and thus understand themselves to be answering a variety of questions as they reply to this 'one'. This variation of understandings cannot but add 'noise' to the data collected thereby. It is, in addition, a fact notable from a number of surveys that a proportion of those who say they do not believe in God nevertheless do not describe themselves as atheists in response to other questions. It may be that they have agnosticism in mind as their preferred self-characterization, but one hypothesis which has gained support is that, even in the largely secular world, there is still a certain social stigma supposed by respondents to be associated with labelling oneself an atheist, a stigma that is not supposed to attach to one as much if one says of oneself merely that one doesn't believe in God. If that is so, the supposed stigma may – one now inclines to hazard – be supposed to attach to one to some extent even if one merely says of oneself that one doesn't believe in God. And, if the last thought hazarded is right, belief in God is therefore likely to be over-reported by surveys. A variant on this worry would be the following: in societies where theism is state-sponsored (e.g. Iran), it seems likely that social/cultural pressures would lead to an over-reporting of theistic belief, even when the anonymity of the respondents was as assured as it ever could be to those who have lived under covert surveillance for so long. Conversely, in societies where naturalism has been state-sponsored (e.g. North Korea), and the same difficulties attend assuring the respondents that they will remain anonymous, theistic belief is likely to be under-reported.

There are a number of variants of worries of the sort outlined in the previous paragraph and I think that one thing one must grant in response to them is that they certainly do point to barriers that lie

between us and knowledge of the prevalence of belief in theism versus agnosticism and atheism. However, another thing one may insist on is that these barriers are not insuperable in principle and, I would further suggest, that they have been overcome in practice by a variety of means, most obviously in the case of the thought hazarded that it might be being supposed by respondents that a 'social stigma' would attach to someone who labelled themselves an 'atheist', one might be able to assure anonymity to the participants (perhaps through making the survey online); one might perhaps misrepresent oneself whilst asking the questions, as for example a member of the National Atheist Association, keen to find out just how many right-thinking people such as oneself there are in the country; and so forth. (Of course the last ruse might lead to an *under*-reporting of theism, but one could compare and contrast the results of surveys done with the ruse and ones done without.) With regard to the worries one might have about respondents from Iran and North Korea, they obviously do not generalize to other countries, or at least they do not apply to other countries to the same extent; and indeed the worries can be mitigated to some extent even for those countries for which they are most pressing; for example, one may ask ex-patriots what they believed when they were in the country in question (although of course care must be taken to control for the fact that they may have been caused to become *ex*-patriots precisely because they had a religious/irreligious view that they feared would lead to persecution if they stayed in their home country). In short, drawing on a variety of surveys is the best way to iron-out problems of this sort, problems which may indeed affect any one of them.

My main worry as I started reading the questions posed by surveys on this issue was that people who said that they believed that there was a God (and, even more so, people who said that they believed that there was not a God) would in general have such confused notions of what sort of God it was they were believing did or didn't exist that there'd be no way of tying their answers to any conclusions about the prevalence of belief in the classical theistic God. I found this worry mitigated, however, by the fact that Gallup Junior and Lindsay, for example, had asked people whether they believed in God or a 'Higher Power'; what, one might think, could be more inclusive than that? A lot of people who aren't theists would therefore have been swept into theistic company in answering that question affirmatively, but, because of this, this survey cannot have over-reported the proportion of respondents who are agnostics and atheists with respect to classical theism. That being the case, I was very surprised to learn that the survey revealed that as many of 39 per cent of British people do not believe in God or a 'Higher

Power'.[3] At least 39 per cent of British people, then, one may confidently infer, are agnostics or atheists with respect to classical theism; were the classical theistic notion of God to be described to them, they would say that they did not believe that anything answered to it.

Finally, it is often suggested that response rates to questionnaires are low enough to render the resulting data statistically questionable and one may reasonably suspect that non-random factors affect who responds to at least some questionnaires. In addition, there can be deliberate attempts to skew the data, motivated by all sorts of purposes. There was recently a campaign – not, one suspects, directed by a desire for truthful reporting – to encourage people, when filling out the 2001 British census, to list 'Jedi' as their religion. In the end a – to me staggering – number of people, 390,127, did state that they were Jedi on the form, *prima facie* suggesting that there were more Jedi in the UK at the turn of the millennium than there were Sikhs, Buddhists or Jews. According to the same census, a full two per cent of the population of my home city, Oxford, were Jedi in 2001.[4] However, in twenty years of walking around the city (even walking on an 'inter-faith' walk down the multi-cultural Cowley Road) I've never seen *any* Jedi at all, or at least seen any that I recognized as such from their appearing in the apparel that the Star Wars films would lead me to think is *de rigueur* for them. What then to say of this batch of worries?

Again, it seems to me that one must concede that there are problems of the sorts outlined in the previous paragraph. But again it seems to me that these problems are not entirely insuperable. There are well-known mathematical techniques for assessing whether or not results are statistically significant and for controlling for at least some of those non-random factors which it has occurred to one may affect response rates. For example, the filling out and returning of the British census is a legal requirement, something which reduces – even if it does not entirely eliminate – grounds for worries of this sort. However, it must be admitted that campaigns aimed precisely at misleading those conducting surveys can remain a step ahead of techniques to avoid being misled. That is, I cannot but believe, what happened in the case of what has become known as 'the Jedi census phenomenon' (it was repeated to varying extents around the Anglophone world at about the same time as the 2001 British census). I just cannot believe – despite the data – that on average one in fifty of the people I see as I walk around Oxford's streets is a Jedi.

In short, then, I would say in response to worries of the sorts I have briefly listed above that they give us reasons for caution when drawing conclusions from the results of surveys of contemporary theistic belief

versus agnostic/atheistic belief, but these reasons can be overstated and these reasons, where and when they are most pressing (and admittedly sometimes they are pressing, e.g. in assessing how many people in North Korea really are atheists or how many people in Oxford really are Jedi), do not undermine the confidence we may reasonably place in some surveys and in the insight they give us into some locations. All of that being so, then, I would maintain that we may sketch with a fair degree of confidence the following picture of the contemporary demographics of theism.

III

There are hardly any (less than ten per cent) agnostics or atheists in North and South America (not including Canada), Africa, the Middle East, India and Southeast Asia. There are more (10–50 per cent) in most European countries and in Canada and Australia/New Zealand; Russia, with 30–40 per cent agnostics and atheists, follows the pattern for Europe. In the Nordic countries, the percentage of agnostics and atheists is highest – at 50–70 per cent. To this general pattern, the most striking exceptions are Vietnam and Japan, where there is much more agnosticism and atheism than their geographical positioning would lead one to expect (70 per cent plus in Vietnam, 60 per cent plus in Japan). Also perhaps exceptional is the Republic of Ireland, where there appears to be much more theism than in the rest of Europe. (Personally, I am sceptical that Ireland really is an exception to the European rule, remembering the truth behind the oft-told anecdote of the tourist travelling in Northern Ireland who is cornered in an alley by a gang and asked threateningly, 'Are you protestant or catholic?' Not knowing the allegiance of the gang but knowing simply that to pick the wrong one of these would be his downfall, he replies, 'I'm an atheist.' The questioner comes back instantly with, 'Ah, but are you a protestant atheist or a catholic atheist?')

On this basis then we may assert with a high degree of confidence that theism is currently very unevenly spread across the surface of the globe: speaking always generally and *ceteris paribus,* we may say that people born in Nordic countries or in Vietnam or Japan are unlikely to become theists; people born in Europe generally (not including the Nordic countries) are slightly more likely to become theists than not; and people born into the Americas (not including Canada), Africa and the Middle East are very likely to become theists. In what follows I shall therefore take these to be established empirical facts. I shall not question

them further. My question is what implications, if any, do these facts have for the rational defensibility of classical theism?

IV

I am not aware of anyone who has advanced an argument to the effect that facts of this sort have positive implications for classical theism; I am aware of two lines of argument that have been advanced for thinking that they have negative implications. I shall describe these two lines of argument, although in this paper I shall only deal in detail with the second, having dealt with the first elsewhere.[5]

First, it is sometimes argued that these facts present a direct problem for *knowledge* of theism, in particular they give one reason to suppose that one cannot know that theism is true even if theism is true, because they show that the facts which lead to belief in theism are too distant from and uncorrelated with the facts that make theism true, too distant and uncorrelated for the belief in theism that results from them to count as an item of knowledge. I will not repeat my counterarguments to such suggestions here, just their conclusion. On the assumption that the empirical facts are as I've sketched them and that theism is true, it follows that – again, it should be stressed, only speaking 'in general' and *ceteris paribus* – people born in Vietnam have a lesser chance of coming to know that it's true than those born in Great Britain, who in turn have a lesser chance than those born in the United States of America. However, that does *not* imply that those who do come to know it don't know it after all: the 'higher-level' chanciness (if indeed that's the best way of thinking of it) affecting whether or not one's a knower doesn't generate a 'lower-order' chanciness in oneself, one that disables one from being a knower. We can see this most quickly by noticing that all sorts of beliefs which obviously and uncontroversially constitute knowledge for some of the people who have them have similar demographics. Access to higher education is very variable across the surface of the globe and a certain higher education is a necessary condition of one's coming to know, for example, that proof of Tychonoff's theorem requires the Axiom of Choice. Suppose one happens to have had such an education and now knows this. One may say that it's a matter of chance – mere good luck – that one was born where one was, had access to the higher education that one did, and thus came to know this. But even so, that chance or luck, one may insist, does not cast into any doubt the process by which one has come to know this, making the process which one's actually employed unreliable (in a way that should trouble an externalist

about knowledge) or the reasons one actually has for believing it any less reasonable (in a way that should trouble an internalist). So much, then, for arguments which seek to use the demographics of theism to undermine it 'de jure', as it were, as an object of knowledge, rather than undermine it 'de facto', as it were, by giving us reason to think it's not true. I shall now turn, then, to arguments of the second sort.

V

Second, then, it is sometimes argued that the facts which I have sketched undermine theism directly as they constitute facts from which one may run a variant of the Argument from Evil against the existence of God: belief in God is, they reveal, distributed other than one would expect were God to exist.

Recently, such an argument has been pressed with much vigour by Stephen Maitzen.[6] It is his argument with which I shall most closely engage in what follows, though the same points may be made against the similar arguments of others. Maitzen suggests that, as a variant of the Argument from Evil, the argument from the demographics of theism has at least this advantage over those versions of the Argument from Evil which focus on the evil of suffering:

> The key difference, I suggest, between suffering and non-belief in God is that suffering is far more evenly distributed than non-belief . . . non-belief in God is anything but uniformly distributed worldwide [and indeed, one could add, over the course of human history], and consequently any explanation in terms of features, such as human freedom, that are uniformly distributed will not work.[7]

But what the argument from the demographics of theism gives with one hand, it threatens to take away with the other. For lack of belief in God, whilst perhaps much more obviously unevenly distributed than the sort of suffering Maitzen has in mind – phenomenological pain – is also much less obviously an evil than that sort of suffering; that sort of suffering is much more obviously an evil, even if it is also perhaps less obviously unevenly distributed. For the argument to give with one hand whilst not taking back in this way with the other, lack of belief in God needs to be linked to something that *is* obviously an evil (ideally, to preserve a comparative advantage, more obviously evil and more evil than phenomenological pain). What could that be?

It is instructive I think that Maitzen narrows his focus to considering a variant of theism that links belief in God with the Highest Good, we may call it 'salvation', by making belief in God a necessary condition of salvation; the variant he has in mind he calls 'evangelical Christianity', though one presumes he would wish to insist that the same points could be made, *mutatis mutandis*, for at least certain Islamic views in addition.[8] Maitzen says this:

> Recall that the problem . . . arises in the first place from the nature of theism's personal creator God, whose perfections include unsurpassable lovingness and who, according to evangelical Christianity anyway, wants everyone to believe the gospel message. Non-belief becomes puzzling if a being of *that* description exists. It therefore does not refute ADH [the Argument from Divine Hiddenness] to construe 'theism' more broadly, as, say, the generic belief in the supernatural: it merely changes the subject.[9]

This is somewhat misleading, on a number of fronts.

First, it ignores what we might call the 'Calvinist' and 'Molinist' explanations of the demographics of theism. I don't favour either of these two solutions, for it seems to me that they don't get rid of the problem of unfairness – they just shift it to an earlier stage in the divine economy; it's no longer unfair that some are born without any chance of coming to the beliefs that would save them because they weren't elected to be amongst the saved anyway (Calvinism) or they had such 'stubborn essences' that God knew that, were he to create them, there'd be no situation in which he could place them in which they'd come to be saved anyway (Molinism). Isn't that just to make the later distribution of belief not unfair by having the relevant unfairness occur earlier, in not electing those souls or in instantiating such stubborn essences? And, in the latter case, could there really be an essence so stubborn that *nothing* God could do would bring the person with it to the relevant beliefs in Him? Be all that as it may, these explanations of the demographics of theism certainly deserve at least a brief mention.[10] Here's the Calvinist one.

Determinism is true and souls may be divided into two kinds, the Elect and the Non-Elect. The Elects' souls are born into societal situations in which they are determined to come to the *ante-mortem* beliefs necessary for salvation; the non-Elects' are born into societal situations in which they're not, indeed they're determined to fail to come to these beliefs. ('Double pre-destination', as it's sometimes called, seems to me

the only way of ultimately making this view coherent.) Calvinism of this sort commits one to what I have elsewhere called 'rather heavy-duty metaphysics' – determinism and souls – but it is not a metaphysics which no theist has ever adopted.[11] Maitzen admits that his 'objection assumes that individuals with one of these characteristics [the previous sentence has included as a characteristic 'one's predestination as non-elect] do not cluster by country or culture'.[12] Indeed it does, but on Calvinism that assumption is obviously false (the demographics show it to be so!), so the objection to theism from its demographics fails. As Jason Marsh has recently done better than I could do in sketching the Molinist solution, I'll simply refer readers to him for that one.[13] I want to move past this family of 'solutions' though and see what can be done if we grant that God does want everyone he creates to believe the gospel message and hasn't created anyone pre-elected not to do so or incapable of actually doing so. If we can come up with a solution within these parameters, we will then be able to preserve the truth of the counterfactual, which common sense suggests is true, namely, a goodly proportion of those who die atheists in Vietnam would have died evangelical Christians in the USA had they escaped from Vietnam to the USA when young enough.

VI

Supposing, then, that we grant that God does want everyone to believe the gospel message and doesn't create anyone who is pre-elected not to do so or couldn't actually do so, we can still resist the sub-conclusion that we should then expect belief in Him to be uniformly distributed over the surface of the globe. And we can resist it in a number of ways. First, the argument depends on the strength we posit that particular divine want to have; more precisely, it depends on our positing that it is strong enough so that it is not over-balanced by other wants for things that are incompatible with it. If we do not grant this, the argument fails. To see why, consider the following by way of analogy.

I want everyone on the planet to buy my latest book,[14] but I also want everyone who does buy it to buy it freely. Freedom being as it is – how Libertarians construe it (see my book for more details) – I cannot then, as a matter of metaphysical necessity, determine things so that the world meets both conditions, so that everyone buys my book *and* everyone who buys it buys it freely. Assuming my power is unlimited, so I could make the world meet either one of the conditions, I'd have to choose which I cared about most before deciding which way to exercise

my power. Let's suppose that I care more about the world being one that meets the latter condition – that everyone who does buy my book buys it freely – than about it meeting the former – that everyone buys my book. If so, you'd expect purchasers to be unevenly distributed over the surface of the globe, in that it would be very unlikely that everyone would freely buy my book and one of the things that leads to people buying a book is the effect on them of other people who've bought it and chosen to recommend it on to their friends; the culture of the society in which they have been brought up; and so forth. Purchasers and non-purchasers would be expected to 'cluster' then over the surface of the globe. Similarly, then, God might well want everyone to believe the gospel message, but he might well want those of them who do so to come to do so without His having had to interfere with libertarian free will; and, if his want for the world to satisfy the latter condition is greater than his want for everyone to believe the gospel message, then, given that one of the things people can use their free will to do is either spread or diminish the spread of the gospel, given that people's receptivity to the gospel is dependent on the culture in which they are brought up and so forth, we'd expect to find that theism had an uneven distribution over the surface of the globe.[15] This is not, note, to suppose that belief is ever under the control of the will, just that spreading or restricting the spread of beliefs sometimes is.[16]

Now this move cannot be made so straightforwardly if one makes having belief in God a necessary condition of the *highest* good, let's call it 'salvation', for then there's nothing God would want more than it on the 'whoever wills the end wills the means' principle. And that, it must be confessed, is what evangelical Christianity, indeed perhaps Christianity more broadly, tends to do: no less an authority than the Athanasian creed, for example, characterizes its statements as those which 'except a man believe faithfully, he cannot be saved'. So, we are most charitable to Maitzen if we think of him as supposing a view of this sort in the background of his argument. It is safe to suppose that God has an *overwhelming* want that everyone believe the gospel message because only by believing it can they be saved – the highest good.

However, with this bit of charity to the argument, we reveal the particularity of the view to which the argument is cogent. It is not an argument against theism per se; it's an argument against this variant of it. Furthermore, we should note that it's an argument only against a variant of this variant, the variant of the variant that makes *ante-mortem* belief in God a necessary condition of salvation. I turn now to consider this point.

VII

Even if I had the power to make everyone buy a copy of my book and I wanted this with such strength that I had no other greater want that over-balanced it (for example, the want that anyone who did buy it would have bought it freely), then one could take the existence of non-purchasers as puzzling only if one supposed that, *in addition to all this*, I had set myself a deadline before which I was to have reached the end I most wanted (everyone having purchased my book) and one supposed that the deadline I had thus imposed on myself had passed. Otherwise, one may simply think of particular non-purchasers or clusters of non-purchasers: 'well, he's not got round to them yet'. There are two plausible suggestions for 'deadlines' for God: either that he has set for each individual his or her own deadline and set it as the moment of his or her death, or that he has set us collectively one and the same deadline, the Eschaton/Last Judgement. An alternative, 'The Harrowing of Hell' model, suggests itself. This was a traditional answer to the question of what happened to those who never heard the gospel; they get a chance when Jesus goes and preaches to them in Hell. Of these three models, only on the *ante-mortem* deadline model does the argument become pressing, because only on this model do the demographics of theism (and – I suppose one must posit additionally – the lack of universality of the phenomenon of deathbed conversions to the right religion) make it that God's missing his deadline. So, *contra* Maitzen, many variants of theism – ones which don't just collapse into the vague belief in the 'supernatural' that he suggests as the (apparently only) alternative to evangelical Christianity – evade his argument. In fact, many variants of evangelical Christianity evade it; it is only variants of theism which make *ante-mortem* belief a necessary condition of salvation that are troubled by it. Even those believers who start off believing in the variant of theism that makes them susceptible to this argument should not then take this argument as giving them overall reason to reject theism unless they think it is more plausible that 'if theism is true, then the troubling variant of it is true' than they think that theism is true. But I take it that it will be much more obvious to everyone who believes that God exists that God exists than that *if* he exists, he's arranged the economy of salvation to depend in this way on people's success in *ante-mortem* theological enterprise, in which case, the argument from the demographics of theism is one that just threatens this peripheral feature of one branch of theological thought; the argument gives nobody reason even to abandon evangelical Christianity.[17]

So, to sum up so far, the demographics of theism *do* give those of us who are theists reason to believe something about our theism. The demographics do not give us reason to think that we don't or probably don't know theism to be true;[18] nor do they give us reason to think that it's false. They do, however, give us reason to think that our believing theism to be true *ante-mortem* isn't as good for us as some have maintained; it may yet be very good indeed for us to believe it, but it's not essential to our ultimate salvation that we believe it. To some evangelical Christians, that may seem a 'concession too far'. They will then naturally find themselves drawn to Calvinism or Molinism, which, it will be recalled, I do not myself favour. Finally then, I wish to do something to show that the 'concession too far' worry is misguided; evangelical Christianity need not be threatened by giving up the claim that *ante-mortem* belief in it is essential to salvation; indeed it is *helped* by giving it up.

VIII

I shall do this in two interlocking ways. First, I'll draw on passages from the Bible – most specifically passages from the New Testament – to suggest what the Christian biblical view is and that it does not suppose an *ante-mortem* belief condition for salvation. I'll argue this through establishing that the New Testament speaks consistently in favour of universalism, which – given the demographics of theism (and the lack of universality of deathbed conversions) – entails the *ante-mortem* belief condition on salvation doesn't obtain. In order to avoid needless misunderstanding at the outset, let me state that Universalism of this sort is quite consistent with only those who accept Christ getting into Heaven and with *post-mortem* punishment for at least some; the only thing it's not consistent with is *post-mortem* annihilation or everlasting Hell for any. Second, I'll look at whether the Christian biblical answer – as I'll by then have established it to my satisfaction – is something which rational reflection on the nature of goodness alone, unaided by revelation, would have led us to expect of God. And I'll conclude that it is *exactly* what we would have been led to expect, the second line of my argument thus giving one a reason to affirm that this is indeed what the Bible should be interpreted as teaching if we wish rationally to believe in it as a revelation from God. Together, then, these two lines of thinking should motivate a rejection of Calvinism and Molinism and, in the light of the demographics of theism (and the absence of universal deathbed conversions), a rejection of the *ante-mortem* belief condition on salvation.

Here's a famous passage from the Bible: 'For as in Adam all die, even so in Christ shall all be made alive.'[19] The clear statement here is that *all* shall be made alive in Christ; the passage doesn't end with 'even so in Christ shall *some* be made alive'. Of course, one might suggest that whilst all are to be made alive in Christ, some of those to be made alive in him are to be made alive merely to be made dead to Him immediately thereafter. But it would be more natural from the surrounding context and other passages to consider being made alive in Christ as being saved. Thus it would be more natural to think that our all being made alive in Christ is our all being raised, being raised to a form of judgement assuredly (there are certainly other passages which suggest judgement), but raised to a form of judgement after which an everlasting life in perfect communion with God awaits. This is re-enforced by other New Testament passages. 'Then as one man's trespass led to condemnation for all men, so one man's act of righteousness leads to acquittal and life for *all* men.'[20] 'For God has imprisoned all in disobedience so that he may be merciful to *all*.'[21] We shall *all* be made alive in Christ; we shall *all* be acquitted and given life. God will be merciful to *all*. One word seems to be popping up quite a bit, doesn't it?[22] Let's look at a couple of other passages.

Romans 10:9 tells us this: 'If you confess with your mouth, "Jesus is Lord," and believe in your heart that God raised him from the dead, you will be saved.' It is certainly true that not all people confess this with their lips or believe this in their hearts *ante-mortem*. In Philippians 2:11 we hear of the – presumably *post-mortem* then – day on which every tongue shall confess that Jesus is Lord and every knee bow before him. I suggest that if one puts the last passage from Romans and the Philippians passage together, universalism becomes pretty much inescapable.

The Romans passage tells us that if you (1) confess with your mouth that Jesus is Lord and (2) believe in your heart that God raised him from the dead, you will be saved. Philippians tells us that one day every tongue will confess that Jesus is Lord and every knee bow before him. So, the Philippians passage tells us that condition 1 as made mention of in the Romans passage will on that day be satisfied by everyone; every tongue will confess that Jesus is Lord. The Philippians passage also tells us that on that day every knee shall bow down before him. That's not quite as clear a statement of the satisfaction of condition 2 – everyone believing in their heart that God raised Jesus from the dead – as would have made the case that universalism is the only biblical option airtight, but it's close. It's just possible to render these two passages consistent

without subscribing to universalism as one might suggest that some of these who on that day confess that Jesus is Lord and bow before him will nevertheless not believe in their hearts that God raised this person – this very person whom they're confessing as Lord and bowing down before – from the dead and thereby they'll nevertheless fail to be saved, failing on condition 2 as stated in the Romans passage. But to suggest that seems, frankly, desperate. Remember they're bowing before him and confessing him as Lord at this stage. Is it really plausible to suppose that some who are doing all this are really still doubting that God raised him from the dead? No, it is not. Thus universalism is – I suggest – pretty much inescapable for anyone who takes the Bible seriously.

I shall now move on to my second line of argument and ask whether universalism is what reason alone would have led us to expect of a perfectly good God.

IX

Obviously it would be good for everyone to enjoy everlasting life with God, and this in itself would give God reason to be a universalist. But there is another thing that is good for people, that others respect their freedom of choice, and this opens up the possibility for wondering, might it be overall good for God to respect our freedom if we freely chose at the Last Judgement to reject His offer of eternal life through Christ and go instead to annihilation or torment in an everlasting Hell? Here I must be brief: it seems to me the answer is 'No'; however good it is to respect free will, it's *ex hypothesi* not as great or good as salvation and thus if God had to choose between them, he'd go for the latter.

Some people have done things which are so terrible that they deserve terrible punishment, punishment which they did not receive this side of the grave. Adolf Hitler and Joseph Stalin would be two obvious examples. Shouldn't God punish them? The answer is obviously 'Yes', but we cannot think that even these monsters committed such terrible crimes that only a punishment of infinite duration could be appropriate and, I suggest, the punishment inherent in their simply being fully in God's presence at the Last Judgement would in itself be all that justice could demand.

Consider Jesus' parable of the returning prodigal. As with any repentance, the destitution and humiliation of the son at the moment of his decision to turn back towards his father is in exact proportion to the selfishness and vanity in which he has previously indulged and we may be assured – if not re-assured – that the same fearful equation will operate

on us in the searing furnace of self-knowledge that must accompany any last judgement. When we are exposed directly to God's glorious presence, the worse we are, the more hellish that refiner's fire will seem to us as it burns off our misplaced egotism and self-satisfaction. It is a fearful thing to fall into the hands of the Living God. There will be plenty of wailing, weeping and gnashing of teeth then – enough to validate the various biblical passages that make mention of these far-from-cheering things. Some have already turned towards God and for them His judgement may seem a momentary delight – the first words of His that they hear will be 'Well done good and faithful servant.' Others will not be turned until they find themselves before Him and to them God will need to say more. To some, the painful truths He must speak on that day will make this judgement seem close to a torturous eternity. But for each of us – from Mother Teresa to Adolf Hitler – the Judgement will in fact pass, once it has done its irreversible perfecting work. And everlasting bliss awaits each of us, once perfected, on the other side.

So, in summary, reflection on the nature of perfect goodness alone would lead us to conclude that which I've argued the New Testament asserts: universalism. That then gives the evangelical Christian reason to reject Calvinism and Molinism as explanations of the demographics of theism and to reject the *ante-mortem* belief condition on salvation.

X

In conclusion, we who are theists must see the demographics of theism not as reason to think that we don't know that God exists, let alone reason to think that God doesn't exist. Rather, we must see these facts as giving us reason to believe what the Bible and natural reason reflecting on the nature of perfect goodness were already giving us reason to believe: the God who exists is one who will not suffer any of us finally to be lost and thus has not made our *ante-mortem* theological success a necessary condition of our being saved. The demographics of theism are thus good news for theists in one sense and they're good news for agnostics and atheists in another (in that they reveal [though not of course to them] that their failure to be *ante-mortem* theists won't ultimately exclude them from salvation). *That* is the gospel message, and it's obvious from the nature of the gospel message (that it doesn't require itself to be believed in order to be true and to be the good news that it is) why it is that God isn't prioritizing making everyone believe it this side of the grave.[23]

Notes

1. I am very grateful to John Cottingham, Douglas Hedley, Dave Leal, Steve Maitzen, Jason Marsh, Yujin Nagasawa and Mark Wynn for their comments on a draft of this chapter.
2. The sources I draw from are mainly those mentioned in the bibliography.
3. Gallup and Lindsay, 1999, p. 121.
4. The relevant data are obtainable from the Office of National Statistics website, at www.statistics.gov.uk/census2001/profiles/rank/jedi.asp/.
5. Mawson, 2009. There's another argument that might suggest itself, springing from the epistemology of disagreement literature: the unevenness of belief in theism suggests that there is persistent disagreement on theism between epistemic peers and where one has that on a topic, one has reason to suppose that knowledge of the topic is not to be had. I won't engage with that argument either.
6. Maitzen, 2006.
7. Maitzen, 2006, pp. 188–9. In fact Maitzen is wrong in this; we'll explore why in more detail in the main text in a moment, but one quick way of showing it is the following: one of the things people can use their freedom to do is either spread or hinder the spreading of belief in God. That being so, free will alone could be used to explain how, over time, we were collectively led from a state in which belief in God was uniformly distributed over the human population (originally just Adam and Eve) to one in which it had the demographics I have sketched. Positing a historical fall would be the simplest way of thus generating a free-will defence to the variant of the problem of the evil posed by the demographics of contemporary theism.
8. Here and elsewhere I can be seen to be assuming false Schellenberg's position that God would have quite generally applicable and overwhelming reasons to remove at least inculpable atheism this side of the grave. The view that the sort of perfect communion with God enjoyed in the beatific vision (the enjoying of which is of course incompatible with one's being an atheist) is the highest good is, I take it, more or less definitive of theism, but that this highest good is also the good of highest urgency is, I take it, not definitive of theism; nor is it plausible. And, in the context of a Schellenberg/Maitzen argument, it's degree of urgency that counts, not degree of goodness. It seems to me that the elimination of atheism can only be made an urgent good by making its achievement *ante-mortem* the necessary means for this sort of *post-mortem* beatific vision.
9. Maitzen, 2006, p. 179; emphasis in original.
10. In a later paper, written in response to one by Marsh outlining the Molinist reply, Maitzen says as much himself. It is from this interchange that I take the phrase 'stubborn essences'. See Marsh, 2008, and Maitzen, 2008. One criticism of Molinism in the main text, that it just pushes the problem back a stage, is in fact *in*applicable to the variant of Molinism that Marsh propounds – unlike the more 'traditional' one propounded by, say, Craig. Marsh is a universalist, who thinks of the demographics of theism as explained by God's grouping those with *ante-mortem* stubborn essences together in Vietnam, for example, for they've nothing to lose by being so grouped. 'As long as God's final victory . . . is taken to mean that salvation will eventually be

achieved by all I cannot see how any of this would be unfair to any of the individuals [born in Vietnam]' (Marsh, 2008, p. 468). In other words, unlike the traditional Molinist view, there are none born with such stubborn essences that they would be *ante-mortem* and *post-mortem* stubborn. Of course, with this version of Molinism there's still the worry that it doesn't preserve the truth of the counterfactual that many who die atheists having been born and lived in Vietnam would have died Christians had they been moved as children to the USA (see main text).

11. See Mawson, 2009, for further discussion of Calvinist-inspired responses to the demographics of theism.
12. Maitzen, 2006, p. 184.
13. Marsh, 2008.
14. Mawson, 2011.
15. Of course, unless one goes for a historical fall (see earlier note), it will seem plausible to maintain that cultural differences have enhanced a pre-existing uneven distribution, not generated it *ab initio*.
16. Here Maitzen would start to object, maintaining that given that belief isn't under the control of the will in any case, so God could, in principle, zap belief in theism into any given individual without thereby immediately interfering with that individual's free will. But I would stress that if the prior lack of belief in that individual had resulted from the free actions of other individuals, not even God could zap belief in theism into the first individual without thereby interfering with the effective freedom of these other individuals.
17. This is a bit 'rough and ready'; note though that I do say that it doesn't give these theists 'overall reason' to reject theism; I would concede that it gives them reasons in the following manner. If you start out believing ' "A" and "If A, then B" ' is true and I give you reason to think that 'A and B' is false, that is my giving you some reason to think ' "A" and "If A, then B" ' is false, which is also then, one might say, my giving you some reason to think 'A' is false, as A's being false is one of the ways in which it could be false that ' "A" and "If A, then B".' But I've not given you overall reason to think A is false if it is 'much more obvious', as I say in the main text, to you that 'A' is true than it is that 'If A, then B' is true. I've given you *overall* reason to think that 'If A, then B' is false.
18. See Mawson, 2006, for the defence of this.
19. 1 Corinthians 15:22.
20. Romans 5:18.
21. Romans 11:32.
22. I've italicized it just in case it's not obvious which one I have in mind.
23. These considerations lead me to think that a stronger argument for atheism could be constructed if one were to focus on the demographics, not of *ante-mortem* belief in God, but of something that is more plausibly the sort of thing that God would have an overwhelming desire to see evenly distributed *ante-mortem* and which, empirical research suggests, has an uneven distribution *ante-mortem*. Discussion with John Cottingham suggests to me that the best candidate for such a thing would be the ability to lead the morally good life, which ability in turn depends on having epistemic access to the nature of the morally good life.

References

Bruce, S. (2002) *God Is Dead: Secularization in the West* (Oxford: Blackwell).
Finngeir, H. (1998) *Atheism in India* (Mumbai: Indian Secular Society).
Finngeir, H. (2003) *Atheism in the World* (Oslo: Human-Etisk Forbund).
Gallup, G., Jr, and M. Lindsay (1999) *Surveying the Religious Landscape* (Harrisburg, PA: Morehouse).
Greeley, A. (2003) *Religion in Europe at the End of the Second Millennium: A Sociological Profile* (New Brunswick, NJ: Transaction Publishers).
Inglehart, R., M. Basanez, J. Diez-Medrano, L. Halman and R. Luijkx (2004) *Human Beliefs and Values: A Cross-Cultural Sourcebook Based on the 1999–2002 Value Surveys* (Mexico City: Siglo Veintiuno Editores).
Maitzen, S. (2006) 'Divine Hiddenness and the Demographics of Theism', *Religious Studies* 42: 177–91.
Maitzen, S. (2008) 'Does Molinism Explain the Demographics of Theism?', *Religious Studies* 44: 473–7.
Marsh, J. (2008) 'Do the Demographics of Theistic Belief Disconfirm Theism? A Reply to Maitzen', *Religious Studies* 44: 465–71.
Mawson, T. J. (2009) 'Mill's Argument against Religious Knowledge', *Religious Studies* 45: 417–34.
Mawson, T. J. (2011) *Free Will: A Guide for the Perplexed* (New York: Continuum).
National Statistics Office, The (2001) *The 2001 Census Online*; available online at www.statistics.gov.uk/census2001/demographic_uk.asp.
O'Brien, J., and M. Palmer (1993) *The State of Religion Atlas* (New York: Simon & Schuster).
Zuckerman, P. (2003) *Invitation to the Sociology of Religion* (London: Routledge).
Zuckerman, P. (2007) 'Atheism: Contemporary Rates and Patterns', in M. Martin (ed.), *The Cambridge Companion to Atheism* (Cambridge: Cambridge University Press), pp. 47–68.

This page intentionally left blank

Part V

Religious Tolerance and Disagreement

This page intentionally left blank

9

Coercion, Consequence and Salvation

Steve Clarke

> Thus Augustine says to the Count Boniface: 'What do these people mean by crying out continually: "We may believe or not believe just as we choose." Whom did Christ compel? They should remember that Christ at first compelled Paul and afterwards taught him.'
>
> – Aquinas, *Summa Theologiae*, 2nd part of the 2nd part, Question 10, Article 8

1. Salvific Exclusivism

Salvific exclusivists believe that there are necessary conditions that must be met before salvation can be attained. Different salvific exclusivists believe in different necessary conditions. Common necessary conditions include: belief in the cardinal tenets of a particular religion, membership of a particular religious organization, conduct of particular religious practices and the avoidance of other practices. Salvific exclusivism stands in contrast with salvific pluralism. Salvific pluralists such as Himma (2002) hold that there is no set of conditions necessary for salvation. On this view, members of many religions are eligible for salvation, and their actual salvation depends on God's consideration of their individual merits. A middle position between these two extremes is one that might be referred to as 'salvific preferentialism'. This is the view that, although God favours those who hold certain religious beliefs, conduct certain religious practices, or are members of particular religious organizations; when deciding whom to grant salvation to, God does not apply hard and fast rules, and will consider the individual merits of those who lack the beliefs, practices and/or organizational membership required for preferential consideration.

Practically all of those who believe in salvation agree that salvation is supremely important. Salvation is typically understood as a precondition to entry into Heaven and the opportunity to enter Heaven is usually understood as an opportunity that only comes along once. On standard Christian and Muslim views entry into Heaven enables the experience of maximal happiness for eternity. And if that were not enough motivation to do what is necessary for salvation, Christian and Muslim theologians often add that some or all of those who are not saved will spend an eternity suffering in Hell. Salvation, if indeed it is available, is of overwhelming importance when compared to any other prudential or 'self-regarding good' (as distinct from 'other-regarding goods'), or indeed all other prudential goods combined. Salvation is a necessary precondition to receiving eternal happiness. In the absence of salvation, our enjoyment of all other prudential goods is necessarily ephemeral.

Salvific exclusivism has long been a prominent position amongst both Christians and Muslims. The Catholic Church officially advocated a fairly strict form of salvific exclusivism up until the time of Vatican II (1962–65), holding that salvation is only available to practicing Catholics (Avalos, 2005, pp. 195–6).[1] Nowadays its position is a salvific preferentialist one, according to which no one is excluded from the bare possibility of salvation; however, membership of the Catholic Church is said to make it much easier to attain salvation (Jones, 1967). The Southern Baptist Convention holds that only Christians can attain salvation, a form of salvific exclusivism that is common amongst conservative protestant groups.[2] Muslim Salvific Exclusivists are often less exclusive than their Christian counterparts. An influential view in Islam is that Salvation is available to 'people of the book' – including Christians and Jews, but not to Hindus, Confucians and others. However, members of the Salafi branch of Sunni Islam are more exclusive than most Muslims, often holding that only very devout Muslims can be saved (Adraoui, 2008).

Consequentialists, who view morality as being exhausted by the consideration of consequence, hold that we ought to do whatever can be done to ensure that the best possible consequences occur.[3] So it seems that consequentialists who are salvific exclusivists should do everything they can to attempt to persuade all others to join the appropriate religious organization, believe the required religious propositions and/or participate in the required religious practices, so that they may maximize their chances of attaining salvation – the one overwhelmingly significant consequence. The same claim can be made for those salvific exclusivists who are not consequentialists, but who accept

that the consequences of salvation are important enough to trump whatever concerns, other than concerns of consequence, that they consider that morality requires. Such concerns might include deontological constraints that would override many considerations of consequence but are not considered to override consequences as significant as the attainment of salvation.[4]

What if the salvific exclusivist who is committed to maximizing the chances that others are able to attain salvation is unable to persuade others to do what she believes that they need to do in order to attain salvation? Sometimes it may be within the power of consequentialist salvific exclusivists to coerce others to join religious organizations, to participate in religious practices, and, in so far as this is possible, to believe particular religious propositions.[5] Should salvific exclusivists who are able to coerce others to do these things, in order to ensure that those others are eligible to receive salvation, do so? Many Kantians would object to the use of coercion, under such circumstances, on the grounds that it fails to respect individual autonomy. But consequentialists should have no such qualms, particularly when the stakes are as high as they are when the possibility of salvation hangs in the balance. Surely the consequentialist salvific exclusivist will consider that she has a moral obligation to employ coercive means to compel acceptance of the correct religion. Furthermore, it seems that the consequentialist salvific exclusivist has a moral obligation to ensure that rival religious doctrines are not promulgated, if there is some chance that these will be accepted by some people and that this acceptance will result in those people being denied salvation.

The last point above is unoriginal. David Lewis (1989) argued, along similar lines, that a devout defender of an orthodox salvific exclusivist religion who was also a consequentialist would have a compelling reason to suppress heresy, which is that it threatens to cause followers of the orthodox religion, who might be susceptible to the influence of heretical teachings, to lose out on all possibility of obtaining salvation. Craig Duncan (2007) extends this argument. According to him, in cases where there is even a small probability that orthodox believers may attain salvation and there is no probability that those who are not orthodox believers will be saved, consequentialism 'absolutely requires' the persecution of unorthodox religion (2007, p. 4).

In reasoning along these lines both Lewis (1989) and Duncan (2007) follow in the footsteps of Aquinas who argued that

> With regard to heretics . . . there is the sin, whereby they deserve not only to be separated from the Church by excommunication, but also

> to be severed from the world by death. For it is a much graver matter to corrupt the faith which quickens the soul, than to forge money, which supports temporal life. Wherefore if forgers of money and other evil-doers are forthwith condemned to death by the secular authority, much more reason is there for heretics, as soon as they are convicted of heresy, to be not only excommunicated but even put to death. (*Summa Theologiae*, 2nd part of the 2nd part, Question 11, Article 3)

Aquinas is no consequentialist, however his argument for harsh treatment for heretics can be understood as a piece of salvific exclusivist reasoning that appeals to considerations of consequence. Heretics have a tendency to promulgate views that would, if accepted, prevent the attainment of salvation. Therefore, considerations of consequence justify the suppression of heretical teachings and justify harsh punitive measures against heresy so as to deter others from promulgating heretical views. Aquinas is far from unique in arguing along these lines. Calvin defended the execution of heretics, as did his successor in Geneva, Beza.[6]

Although Aquinas's views made appeal to considerations of consequence in justifying the suppression of heresies, he did not extend this argument as far as it might go. He did not advocate the general suppression of other religions or the use of forced conversion for the sake of salvation. Indeed, Aquinas upholds the standard Christian line of his day, in insisting that belief must be voluntary (*Summa Theologiae*, 2nd part of the 2nd part, Question 10, Article 8) but arguing that heretics and apostates constitute a special case, because their previous expressions of faith amount to a promise which they can be held to, and which cannot be renounced.

This standard line was by no means universal amongst Christian theologians. Duns Scotus notoriously argued for the forced baptism of Jews. His argument turns on the Orwellian sounding doctrine of *consenting virtually*, according to which assent to baptism due to fear of injury or death can be sufficient to count as genuine consent in the eyes of God (Turner, 2006, p. 196). Of course, once a legitimate baptism had occurred, Jews were no more able to leave the Church than heretics or apostates. Duns Scotus supplements this argument with an argument for the compulsory baptism of Jewish children that rests on a conception of parental rights which broke with the Christian orthodoxy of the day. Whereas Aquinas and others had declared that one should not forcibly baptize the children of Jews and heathens, on the ground that this would be a violation of the rights of their parents, Duns Scotus argued

that parental rights should give way to the rights of higher powers. According to him the rulers of Christian states, who were intermediate between ordinary folk and God, have a duty to enforce the rights of God, the highest power, to have Jewish children converted to Christianity (Krop, 1989, pp. 164–5). If their parents were to resist this then their resistance could trigger an additional argument for their forcible baptism. According to Duns Scotus:

> it is religiously just for those parents themselves to receive baptism forcibly with threats and fear, because although they will not be real believers at heart, the evil for them to be stopped from serving their law with impunity is less than serving that law freely. What is more, if their children are well educated, they will be real believers in the third and fourth generation.[7]

None of the above argument turns on any particular details of Judaism, so it looks like it can be applied equally to all non-Christians.

Lewis (1989) and Duncan (2007) are concerned, *inter alia*, to demonstrate that salvific exclusivist consequentialist arguments for the suppression of unorthodox religions are valid. Here it will be taken as given that such arguments, along with arguments for coerced religious conversion, are valid. Our focus will be on a different problem. Many of us live in liberal societies and religious tolerance is a core value of contemporary liberal societies. We may not share the religious beliefs of the Hare Krishna living at the end of our street and we may not like her religious practices, however, we put up with her presence and try not to express our distaste for her behaviour. In some instances our decision to be tolerant is pragmatic and in some instances it is based on a commitment to the value of religious tolerance. This commitment is common to almost all variants of liberalism; and religious tolerance is often upheld as a core value of contemporary liberal societies.

Not all citizens of liberal societies are convinced that religions other than their own should be tolerated. Prominent amongst the unconvinced are salvific exclusivists. There are many who hold salvific exclusivists religious beliefs and from the salvific exclusivist point of view, as we have seen, there are compelling reasons to oppose the toleration of other religions. Of course not all salvific exclusivists will find such arguments compelling. Some will consider that the values of religious tolerance and/or autonomous choice outweigh the consequence of other people not being eligible for salvation. But many will follow Aquinas and Calvin in finding such arguments extremely compelling.

The problem that will concern us here is the problem of when and how members of a liberal society who value religious tolerance can share their society with those salvific exclusivists who find arguments for religious intolerance that appeal to considerations of consequence compelling. These will include consequentialists who are salvific exclusivists and they will also include those non-consequentialists who decide that whatever deontological constraints on their behaviour there are, that these are not sufficient to override the moral importance of ensuring that others attain salvation, even if this involves using coercive measures. For convenience we will refer to all of those who accept a moral argument for using coercive measures to ensure that others are eligible to attain salvation as 'interventionist salvific exclusivists'.

2. The Interventionist Salvific Exclusivist in a Liberal Society

Interventionist salvific exclusivists are motivated to try to ensure that others are eligible for salvation and so they are motivated to try to ensure that others meet whatever conditions this involves. Most of the time these conditions will include accepting some or all of the key aspects of a particular religion, whether this be acceptance of core beliefs, membership of a particular religious organization, or participation in particular religious practices. I will refer to any or all of these as 'acceptance' of a religion.

Here our concern is with interventionist salvific exclusivists who reside in typical liberal societies and in typical liberal societies individuals enjoy freedom of religion. They are free to choose which religion to accept. They are free to switch allegiance to a different religion at any time, and they are free to proselytize on behalf of their favoured religion. In one way freedom of religion should be welcomed by interventionist salvific exclusivists; it ensures that they are able to attempt to convert everyone within a society to acceptance of their religion. However, there is another way in which freedom of religion is unwelcome for interventionist salvific exclusivists. Precisely because liberal states uphold the value of freedom of religion, they restrict the use of some of the means to make converts that the interventionist salvific exclusivist may wish to employ and which may be most effective in securing conversions. In particular, liberal states restrict the use of coercive means to secure conversions, especially the use of violence and the use of the threat of violence.[8] They do so because the use of coercive means diminishes a

person's freedom of choice, a core value that liberal states traditionally aim to uphold (Spector, 2008).

The benefits of coercion for the interventionist salvific exclusivist are not exhausted by opportunities to make additional conversions. The interventionist salvific exclusivist is in competition with other religions, many of which have an interest in making apostates of the followers of her religion. If believers in these other religions are also interventionist salvific exclusivists, then, all things being equal, they will be as motivated as she is to make conversions. But this is not the only sort of threat she faces. Because, in the liberal state adherents to her religion are not compelled to accept any religion, they may become agnostics or atheists. They may also create new religions. If these bear a partial resemblance to the religion that the interventionist salvific exclusivist seeks to promote then they are heresies from the point of view of the interventionist salvific exclusivist and they constitute a particular danger to the interventionist salvific exclusivist because current adherents only have to change some of their beliefs, while being able to retain others, to convert to the heresy, which is generally easier to do than to convert to an entirely new religion.[9] All of these threats can be countered by the use of coercion.

Although liberal states oppose the use of coercion by some citizens against other citizens, most liberal states typically suffer no qualms about employing coercive means, through laws backed by standing police forces and armies to prevent individuals and groups residing within the state, from using coercive means against one another. Because the legal and institutional apparatus of the liberal state stands in the way of the interventionist salvific exclusivist's goals, interventionist salvific exclusivists will be motivated to attempt to overthrow the liberal state and perhaps to collaborate with other opponents of the liberal state to attempt to overthrow that state. So, the presence within a liberal state of interventionist salvific exclusivists who have a disposition to overthrow the state is a standing threat to that state. If some interventionist salvific exclusivists are considered too much of a threat to a liberal state then that state might feel warranted in suppressing some or all salvific exclusivist religions. It is perhaps unclear whether a society that decided to do this and to not 'tolerate the intolerant' would count as a genuine liberal society or not.[10] But this definitional issue will not be taken up here. From the point of view of tolerant mainstream liberals, a society that tolerated all religions except some salvific exclusivist religions would be far preferable to the society that

the interventionist salvific exclusivist would prefer, in which only one religion was tolerated.

The interventionist salvific exclusivist is motivated to maximize the number of people who adhere to the correct religion and are therefore eligible to attain salvation, and she will feel justified in using coercive means to achieve this goal. However, if she attempts to use coercive means in a liberal state, she can expect to be prosecuted by the state, and may become less able to make converts in the future. Repeated attempts to use coercive means may also result in the practice of her religion being suppressed, in which case many who previously followed her faith may abandon that faith and may become ineligible for salvation as a result. So, the potential gains to be had by the use of coercive means must be great if they are to justify the use of coercive means.

Because the liberal state can be expected to try to ensure that there are no lasting benefits to be gained from the use of coercive means, by some of its citizens against others, the only benefits that can reasonably be expected to be reliably available to the interventionist salvific exclusivist, which might compensate for the risks involved in using coercive means, are ones that may be achieved by overthrowing a liberal state and replacing it with a religious state. If this could be done, then it would enable the interventionist salvific exclusivist to use coercive means in an unhindered manner, to ensure that the great majority of people living in that state meet the necessary conditions of salvation. Furthermore, it would enable the interventionist salvific exclusivist to redeploy the legal and institutional apparatus of the state in aid of this goal. In effect the salvific exclusivist is faced with a basic choice between attempting to 'save souls' via non-coercive means, while accepting that many may be lost, and attempting to overthrow the state in order to create a religious state in which many more souls may be saved. However, the act of attempting to overthrow the state is fraught with danger and may result in the loss of many souls if it fails.

3. Modelling the Interventionist Salvific Exclusivist's Decision

The interventionist salvific exclusivist aims to maximize the number of people who accept the correct religion, so as to maximize the number of people who are eligible to attain salvation. From her point of view all other considerations are trumped by this one consideration. In order best to achieve her goal the interventionist salvific exclusivist who is living in a liberal state must make a basic choice. She can attempt to

operate within the rules of the liberal state or she can involve herself in the organization of an attempt to overthow the state and replace it with a religious state. Either course of action involves potential risks and potential benefits. If she attempts to operate within the rules of the state and proselytizes on behalf of her favoured religion she may make converts and she may lose adherents to competitors. If she attempts to overthrow the state she may succeed and be able to compel acceptance of her favoured religion and she may fail and her cause may be harmed as a result. How is she to think about this choice?

Contemporary work on decision-making under risk provides us with two broad frameworks which may be used to guide her decision. These are cost–benefit analysis (CBA) and the precautionary principle (PP). We will now consider how each may be used to shape her decision.[11] Using CBA a comparison of the expected utility of the two courses of action is made and the course of action with the greatest expected utility is recommended. Suppose, to consider a somewhat overly simplistic example, that 50 per cent of a population of 100,000 who live in a small liberal state are adherents to a particular religion and are considered to be eligible for salvation by an interventionist salvific exclusivist. She calculates that if she and other interventionist salvific exclusivists who follow the same religion devote their efforts to non-coercive conversion over the next ten years, they stand a 75 per cent chance of converting another 10 per cent of the population, or 10,000 souls and a 25 per cent chance of losing 20 per cent of the 50,000 faithful, or 10,000 souls. So she calculates that the expected utility of non-coercive conversion is $(10,000 \times 0.75) - (10,000 \times 0.25) = 5000$. She then calculates that if she and other interventionist salvific exclusivists attempt to overthrow the state they stand a 50 per cent chance of succeeding and a 50 per cent chance of failing. She further calculates that if they succeed they can expect to convert approximately 40 per cent of the current population (she assumes that approximately 10 per cent will flee to another state) or 40,000 souls. However, if they fail they can expect to lose 80 per cent of the faithful, or 40,000 souls (she assumes that 20 per cent will continue to practice their religion even if they are persecuted for doing so after the failed coup). So she calculates that the expected utility of attempting to overthrow the state is $(40,000 \times 0.5) - (40,000 \times 0.5) = 0$. This application of CBA suggests that the true believers should not now resort to coercive means, but the difference in expected utility between the two possible courses of action is not overwhelming, so it seems that she should review her decision at a later date as the factors informing her calculation may well change.[12]

It might be thought that only leaders of religious groups that can count on the support of a significant proportion of the population will stand any serious chance of overthrowing a state, and that therefore we can dismiss the possibility that interventionist salvific exclusivists who apply CBA, and who adhere to a religion that holds sway over a small minority of the population, will attempt to overthrow the state. But this would be to misunderstand the calculation at hand. Suppose that a salvific exclusivist religion was followed by 2 per cent of the population of 100,000 and suppose that their interventionist leader made the calculation that they had a 5 per cent chance of taking over the state and converting the remaining 98 per cent of the population and a 95 per cent chance of failing in their attempt, leading to suppression and the loss of all of the souls that are currently believed to be eligible to be saved. On these calculations the expected utility of attempting a coup is $(0.05 \times 98000) - (0.95 \times 2000) = 3000$. So, it will be rational for the salvific exclusivist leader to attempt a coup under these circumstances, unless the expected utility of making conversions over an equivalent period of time (say ten years) is 3000 or over, which would involve a very high rate of conversion, for a religion that currently only has 2000 followers in a population of 100,000.

A possible objection to the way I have set up these calculations is that if the experience of maximal happiness for eternity has infinite value then the value of the salvation of any number of individuals (above zero) will be infinity; and therefore it is pointless to try to increase the number of people who will attain salvation beyond one. So, saving an additional 3000 souls should make no difference to the interventionist salvific exclusivist, provided that at least one soul is already going to be saved. Intuitively, many of us will want to be able to say that a greater number of people experiencing infinite utility is preferable to a smaller number, despite the problem of comparing infinities. Vallentyne and Kagan (1997) appeal to nonstandard mathematics to defend the coherence of this intuitively appealing assertion.[13] It is beyond the scope of this essay to determine whether they are successful. If they are, then their approach can be used to meet the objection. Another way of meeting the objection is also available, which is to deny that the experience of maximal happiness over an infinite period of time has infinite utility. If the utility of happiness is discounted at an appropriate rate over time, then infinite happiness can be ascribed a very high (but finite) utility (Garcia and Nelson, 1994, p. 185).

Examples like the above, where the use of CBA can lead to the recommendation that we take great risks, has led some to be sceptical of

the appropriateness of CBA as a guide to decision-making under risk.[14] Such sceptics are typically advocates of the PP. Despite being referred to as *the* PP, there is no one PP. In fact, there are at least twenty different formulations of the PP (Sunstein, 2005, p. 18). What they have in common is that they all involve an attempt to capture the apparently commonsensical idea that we are 'better safe than sorry' (Sandin, 2007, p. 105). In other words, we should err on the side of caution even if that means foregoing significant potential benefits. The recommendations resulting from applications of CBA and the PP to the same set of circumstances will be much the same when potential benefits are not significant enough to outweigh risks. However, when potential benefits are significant enough to outweigh risks, and risks are not trivial, then applications of CBA and typical variants of the PP will lead to very different recommendations. While an application of CBA will typically lead to the recommendation of action in such circumstances, applications of most versions of the PP will lead to the recommendation of inaction.

From the point of view of a salvific exclusivist religious leader who is contemplating a coup, in order to be able to compel people to accept the religion that will enable them to attain salvation, there are almost always significant risks involved in action as a failed attempt to conduct a coup may well lead to the suppression of the religion she wishes to promote. It looks like use of the PP would only ever lead to the recommendation of attempting a coup when it was certain that it would be successful or when it was certain that there would be no adverse consequences that would result from a failed coup. But this is effectively never.

Despite the above line of reasoning, I am not convinced that applications of the PP do invariably lead to the recommendation of inaction when an interventionist salvific exclusivist leader is contemplating attempting a coup. To see why we need to consider the following influential line of criticism of core versions of the PP, I refer to Manson (2002), Sunstein (2005) and others. These critics argue that it is impossible to apply these core versions of the PP in a logically consistent way and that actual applications of the PP necessarily involve a failure to recognize that there are 'risks on all sides', and a concomitantly selective approach to the consideration of risks. For example, the PP is often applied to justify restriction of the production of Genetically Modified (GM) crops on the grounds that the consumption of GM foods might possibly be a health risk to humans. But this line of reasoning fails to take into consideration the possibility that a failure to grow GM crops will pose a risk to human health. Non-genetically modified foods are

generally more expensive than GM foods and in situations of famine the worst off face the possibility of being unable to afford proper nutrition if they do not have access to GM crops. If we focus on this risk then the PP seems to lead to the recommendation that we should definitely grow GM crops.[15]

It might seem that an application of the PP would lead the salvific exclusivist to decide not to risk losing the souls that are currently eligible for salvation and therefore to avoid attempting to overthow the state. However, this appearance is only manifested when we lose sight of the risks involved that are borne by the unconverted in being denied the possibility of salvation. A salvific exclusivist who focused on the risks to the unconverted (and did not consider the risks to those that currently adhere to the correct religion) might equally employ the PP to argue that we should avoid the risks to the unconverted involved in being denied salvation, and so attempt to overthrow the state in order to be able to coerce acceptance of the correct religion, even if the chances of success are extremely low.

Manson (2002) and Sunstein (2005) are right to hold that we cannot apply the PP in a logically consistent manner, but this does not mean that we cannot apply the PP at all. We can apply it when we take a selective approach to risk. According to Sunstein selective approaches to risk are widespread and are the result of the operation of a variety of cognitive biases, including the availability heuristic, probability neglect, loss aversion, a belief in the benevolence of nature and systems neglect (2005, p. 35). The most significant of these biases, in Sunstein's view, are the biases accompanying the use of the 'availability heuristic', a rule of thumb which is used pervasively in the lay assessment of risk (Sunstein, 2005, pp. 35–9). Risks will seem highly 'cognitively available' to us when we are familiar with them or when they are made highly salient to us by recent events, by the testimony of others or by the media. If a type of dangerous event has been experienced recently, is discussed widely, or is highlighted in the media then people will intuitively raise their assessment of its likelihood of recurrence and this may 'crowd out' awareness of other risks. After the events of 9/11 the risk of a terrorist attack while flying was highly cognitively available to many people, particularly in America. This caused many people to use alternative forms of transport, such as driving, for inter-city trips that they would have previously flown on. These alternatives involve risks of their own. However, because the cognitive availability of terrorist activity on airplanes crowded out these other risks, many people did not consider them when making inter-city transport choices.[16] The pervasiveness of

cognitive biases in human cognition, and of biases resulting from the use of the availability heuristic in particular, helps explain why the PP can be both incoherent and widely applied (Clarke, 2010).

It is not clear that a committed interventionist salvific exclusivist who was contemplating a coup in order to maximize the number of people who were eligible for salvation would focus on the risks involved in attempting to overthrow the state. She might instead focus on the risks to those who are currently ineligible for salvation. This is, after all, the primary concern of the interventionist salvific exclusivist. So, in her hands, an application of the PP might recommend attempting to overthrow the state, even in circumstances where applications of CBA would lead to the recommendation of inaction.

4. Liberal Responses to the Threat of Salvific Exclusivism

How should a liberal state respond to the possible threat posed by interventionist salvific exclusivists, given that the reasoning of salvific exclusivist leaders who are influenced by considerations of consequence may lead them to attempt to overthow the state? The first thing to say is that the state should attempt to understand the moral reasoning that accompanies particular instances of salvific exclusivism. Some salvific exclusivists will take the view that they should not attempt to coerce others into accepting any particular religion, even if this is in their own interests; in which case they will be unlikely to be motivated to attempt to overthrow the state. However, others may well be swayed by the argument that we have rehearsed and will consider it appropriate to coerce people to accept the religion that offers the only possibility of their salvation.

The second thing to say is that the state should attempt to monitor the activities of salvific exclusivist leaders who are liable to be swayed by interventionist arguments for coercion, as these leaders may be highly motivated to attempt to overthrow the state or to collaborate with others in attempting to overthrow the state.[17] Often states have more than one group of enemies and interventionist salvific exclusivist leaders who are disposed to overthrow the state may ally themselves with other opponents of the state in order to achieve their long-term goals.[18]

The third thing to say is that a liberal state should take all reasonable steps to ensure that when interventionist salvific exclusivist leaders decide whether to attempt to overthrow the state or not, the answer they are most likely to be led to is that they should not attempt to overthrow the state. As we have seen, some interventionist salvific exclusivist

leaders may try to make this decision on the basis of reasoning that is guided by CBA, while others may try to make it on the basis of reasoning that is guided by the PP. We will consider what the liberal state can do to influence leaders guided by each of these broad frameworks.

The government of a liberal state will typically be able to exert an influence on at least three variables that inform the calculation that the interventionist salvific exclusivist leader who applies CBA will try to make when deciding whether or not to attempt to overthrow the state. They can affect (1) the opportunity of the salvific exclusivist to pursue non-coercive means to achieve their goals, (2) the possibility that an attempt to overthrow the state may succeed and (3) the consequences of conducting a failed coup. I will now briefly consider all three of these variables.

Because the interventionist salvific exclusivist will be comparing the consequences of non-coercive attempts to make conversions with the consequences of attempting to overthrow the state, it is vital, from the point of view of the liberal state, that the interventionist salvific exclusivist is free to make conversions. If the interventionist salvific exclusivist is completely unable to make conversions by operating within the rules of the state then any possibility of overthrowing the state will be sufficient to motivate the interventionist salvific exclusivist to attempt to overthrow the state. And if the freedom to make conversions is restricted to some degree then this will reduce the prospect of the interventionist salvific exclusivist making conversions and so make the consequences of an attempt to overthrow the state appear comparatively more appealing. In general, the lower the possibility that the interventionist salvific exclusivist believes the prospects of making conversions (and retaining the loyalty of those who are currently committed to their favoured religion), the more appealing the prospect of attempting to overthrow the state is going to appear.

The liberal state can take a variety of measures to reduce the chance that a coup initiated by interventionist salvific exclusivists is successful. These will include taking reasonable steps to ensure that the loyalties of those individuals who make up the instruments of state power – the police force, the army, the judiciary and so on – are primarily committed to the state rather than to religious organizations. It can also include ensuring that religious organizations are not able to acquire experience in military operations, and are not able to create ongoing associations with paramilitary groups.

When conducting a CBA to decide whether to attempt to overthrow the state or not, the interventionist salvific exclusivist will need to have

in mind the possible consequences of failure. The possible consequences that the interventionist salvific exclusivist will care about most will be the loss of faith by current believers and the possibility of being prevented from converting others to the interventionist salvific exclusivist's religion in the future. Knowing that these possibilities are matters of grave concern to the interventionist salvific exclusivist places the defender of the liberal state on the horns of a dilemma. On the one hand, one of the core values of the liberal state is freedom of religion. On the other hand, if the interventionist salvific exclusivist believes that she and others will still be free to proselytize on behalf of her religion after an attempted coup, then they will be more likely to attempt a coup, so a willingness to restrict the freedom of the interventionist salvific exclusivist under certain conditions makes it more likely that freedom of religion will be available to all. But this is really just one instantiation of a more general dilemma that the defender of liberalism must face in many contexts. Liberals are generally in favour of tolerance but some of those that they may decide to tolerate represent a threat to the future of liberal states. It is beyond the scope of this essay to solve this dilemma. I simply note here that one way that liberals can make it less likely that salvific exclusivists will attempt to overthrow the state is by making it clear to interventionist salvific exclusivists that their freedom to proselytize of behalf of their religion will be significantly restricted if they do attempt to overthrow the state and fail.[19]

Suppose that an interventionist salvific exclusivist leader's decision about whether to attempt to overthrow the state or not is based on the PP rather than CBA. Is there anything that the state can do to reduce the chance that she will attempt to overthrow the state? I will argue that there is. Recall that the interventionist salvific exclusivist leader's decision under an application of the PP depends crucially on which risks are salient to her when she is attempting to apply the PP. If she attends to the risks associated with a failed coup then an application of the PP will lead her not to attempt to overthrow the state. However, if she attends to the risk that people may be ineligible for salvation because they do not accept the required religion then an application of the PP will lead her to attempt to overthrow the state. If she attends to both of these, then, for the reasons that Manson (2002) and Sunstein (2005) point out, she will be unable to apply the PP. However, as we saw earlier, she is unlikely to attend to both sets of risks when either one is highly 'cognitively available' to her.

Cognitive availability is influenced, *inter alia*, by familiarity. If people know about a risk and are reminded of it on a regular basis, then this

risk will be less likely to be 'crowded out' by other risks, and it will be more likely to crowd out other risks. It is heavily in the state's interest to ensure that the risks associated with attempting a coup are familiar to the interventionist salvific exclusivist leader who applies the PP and to any other salvific exclusivists who might become interventionist salvific exclusivist leaders and might apply the PP. If these risks are familiar then they will be cognitively available and will be hard to crowd out. If they are not crowded out then the interventionist salvific exclusivist leader's application of the PP will either lead to the decision not to attempt a coup or will lead to a failure to apply the PP.

The risks of attempting a coup will tend to be familiar when salvific exclusivists have been told about such risks sufficiently often. Such information can be imparted in various ways. For example, the education system could be used to teach people about the punitive response of the state to attempted coups and about past failed coups, attempted coups that have failed could be celebrated (the celebration of Guy Fawkes Night – also known as Bonfire Night – in England is a precedent here) and current plots to overthrow the state that fail could be discussed widely by state representatives (and others) in the media.

5. Concluding Remarks

The combination of acceptance of salvific exclusivism and a commitment to the moral goal of maximizing the best consequences for others can be extremely dangerous from the point of view of liberal societies. In virtue of their intellectual commitments, interventionist salvific exclusivists consider that they have a moral obligation to ensure that everyone accepts their religion. If others will not accept their religion voluntarily, then they have a compelling reason to attempt to coerce them to accept that religion, and a compelling reason to use coercive measures to prevent proselytizing on behalf of other religions. If the state tries to prevent them from using coercive means to achieve these ends then they will have a compelling reason to attempt to overthrow the state.

There are few sorts of states that would be able to tolerate the presence of interventionist salvific exclusivists other than religious states that happen to endorse their religion. However, it is sometimes possible for a liberal state to tolerate the presence of interventionist salvific exclusivists. To do so the liberal state needs to ensure that the interventionist salvific exclusivist will judge that she is more likely

to be able to make converts by non-coercive means than she is by attempting to overthrow the state and thereby enabling the use of coercion.[20]

Notes

1. The Catholic Church has generally made exceptions for morally upstanding people who lived before the birth of Christ.
2. Specifically, Southern Baptists hold that 'There is no salvation apart from personal faith in Jesus Christ as Lord' (see the Current Baptist Faith and Message Statement: www.sbc.net/bfm/bfmcomparison.asp; accessed 9 February 2011).
3. Strictly this claim is only true for 'expected value consequentialists' (Duncan, 2007, pp. 6–7). A further complication is that whereas on most views salvation is available to an unlimited number of people, Jehovah's Witnesses believe that salvation is only available to a limited number of people.
4. Another possibility is that the same argument might be compelling for strong salvific preferentialists, who hold that God is unlikely to grant salvation to very many of those who do not accept the correct set of beliefs, join the right organization or conduct the right practices. But we will set this possibility aside for the sake of simplicity.
5. Is it possible to coerce someone to believe that a particular religious proposition is true? Locke (1991) famously argued that it is not. Waldron presents a counterargument (1991). I will not attempt to judge the issue here.
6. Duncan (2007, pp. 2–4) presents relevant texts in which all three authors argue for the violent suppression of heresy on salvific grounds.
7. Cited in Krop, 1989, p. 165.
8. The exact definition of coercion is a complex exercise and is often said that coercion includes various forms of psychological pressuring, including the use of 'social pressure' and emotional manipulation. Most liberal states are not in a position to prevent these from being used, and may not be opposed to their use in any case. For more on the definition of coercion see Anderson, 2006.
9. This consideration may go much of the way to explaining why Aquinas and other leading Christian theologians have advocated particularly harsh treatment of heretics.
10. For a recent discussion of the limits of tolerance in liberal societies see Ignatieff, 2004.
11. It is sometimes suggested that CBA and the PP should not be directly compared because, while CBA applies to situations of risk (where the probabilities of potential outcomes may be specified), the PP applies to situations of uncertainty (where the probabilities of potential outcomes may not be specified). However, most real world situations involve a mix of risk and uncertainty. In these many cases both CBA and the PP can be applied, so it seems appropriate to compare the two (Clarke, 2009, pp. 161–2).
12. The interventionist salvific exclusivist should also consider the possibility that any new state that is established after a successful coup may fail to continue to deliver the benefits that motivated the coup (in this case the

ongoing opportunity to coerce acceptance of a particular religion). Thanks to Russell Powell for this nice point.

13. Strictly, Vallentyne and Kagan (1997) are defending the coherence of a slightly broader claim than the one with which I am concerned. They are concerned to defend the coherence of comparisons involving infinite numbers of 'locations', which include infinite series of temporal locations (eternity) as well as infinite populations of people.
14. For a defence of CBA against a variety of criticisms see Schmidtz, 2001.
15. There are some versions of the PP that do allow for the consideration of benefits, which are then weighted against risks. But such 'weak' versions of the PP are not importantly distinct from CBA, so they are not part of an alternative approach to decision-making under risk which we need to consider here. For more on the distinction between 'weak' and 'strong' versions of the PP and their relation to CBA see Sunstein (2005, pp. 18–26) and Clarke (2009, pp. 163–5).
16. Gigerenzer (2004) estimates than an extra 350 road fatalities occurred in America in the final three months of 2001 as a result of people avoiding air travel following the events of September 11 2001. This is a higher number of deaths than the 266 deaths involved in the four crashed flights of September 11 2001 combined.
17. In some liberal states, such monitoring may be problematic due to concerns about privacy. It is hard to know how to balance concerns about privacy with concerns about the security of a state and there is a lack of agreement in contemporary liberal states about how to strike such a balance.
18. Even failed attempts to overthrow liberal democratic states can be damaging as liberal democratic states may respond to such attempts by becoming more willing to restrict the liberties of minorities, in order to minimize the possibility of being overthrown. Indeed, liberal democratic states have a long history of reacting in exactly this way in response to real and perceived threats (Ignatieff, 2004, pp. 54–81).
19. The religion in question will, in all likelihood, receive extremely bad publicity in the media in the event of an unsuccessful coup, and this may be a further barrier to future conversions.
20. Thanks to participants in 'The Concept of God and the Cognitive Science of Religion: an International Conference' held at the University of Birmingham in 2009, as well as to audiences at the Centre for Applied Philosophy and Public Ethics, Canberra Division, Oxford University and the University of South Carolina, and also to Russell Powell and Yujin Nagasawa for helpful comments on earlier versions of this essay.

References

Adraoui, M.-A. (2008) 'Purist Salafism in France', *ISIM Review* 21: 12–13.

Anderson, S. (2006) 'Coercion', *Stanford Encyclopedia of Philosophy*; available online at www.seop.leeds.ac.uk/entries/coercion/ (accessed 3 February 2011).

Aquinas, T., *Summa Theologiae*, English trans. (1265–74); available online at www.newadvent.org/summa/index.html (accessed 3 February 2011).

Avalos, H. (2005) *Fighting Words: The Origins of Religious Violence* (Amherst: Prometheus Books).
Clarke, S. (2009) 'New Technologies, Common Sense and the Paradoxical Precautionary Principle', in P. Sollie and M. Duwell (eds), *Evaluating New Technologies: Methodological Problems for the Ethical Assessment of Technology Developments* (Dordrecht: Springer), pp. 159–73.
Clarke, S. (2010) 'Cognitive Bias and the Precautionary Principle: What's Wrong with the Core Argument in Sunstein's *Laws of Fear* and a Way to Fix It', *Journal of Risk Research* 13: 163–74.
Duncan, C. (2007) 'The Persecutor's Wager', *Philosophical Review* 116.1: 1–50.
Garcia, J. L. A., and M. T. Nelson (1994) 'The Problem of Endless Joy: Is Infinite Utility Too Much for Utilitarianism?', *Utilitas* 6: 183–92.
Gigerenzer, G. (2004) 'Dread Risk, September 11 and Fatal Traffic Accidents', *Psychological Science* 15: 286–7.
Himma, K. (2002) 'Finding a High Road: The Moral Case for Salvific Pluralism', *International Journal for the Philosophy of Religion* 52: 1–33.
Ignatieff, M. (2004) *The Lesser Evil: Political Ethics in an Age of Terror* (Edinburgh: Edinburgh University Press).
Jones, E. M. (1967) *The Church-God's Plan for Man: The Teaching of the Second Vatican Council* (London: Burn and Oates).
Krop, H. A. (1989) 'Duns Scotus and the Jews: Scholastic Theology and Enforced Conversion in the Thirteenth Century', *Nederlands Archief voor Kerkgeschiedenis* 69: 161–75.
Lewis, D. (1989) 'Mill and Milquetoast', *Australasian Journal of Philosophy* 67.2: 152–71.
Locke, J. (1991) 'A Letter concerning Toleration', in J. Horton and S. Mendus (eds), *John Locke: A Letter concerning Toleration in Focus* (London: Routledge), pp. 12–56.
Manson, N. A. (2002) 'Formulating the Precautionary Principle', *Environmental Ethics* 24: 263–74.
Mill, J. S. (1859) *On Liberty* (London: J. W. Parker and Son).
Sandin, P. (2007) 'Common Sense Precaution and Varieties of the Precautionary Principle', in Tim Lewens (ed.), *Risk: Philosophical Perspectives* (London: Routledge), pp. 99–112.
Schmidtz, D. (2001) 'A Place for Cost-Benefit Analysis', *Philosophical Issues (A Supplement to Nous)* 11: 148–71.
Spector, H. (2008) *Autonomy and Rights: The Moral Foundations of Liberalism* (Oxford: Oxford University Press).
Sunstein, C. R. (2005) *Laws of Fear: Beyond the Precautionary Principle* (Cambridge: Cambridge University Press).
Turner, N. L. (2006) 'Jewish Witness, Forced Conversion, and Island Living: John Duns Scotus on Jews and Judaism', in Michael Frassetto (ed.), *Christian Attitudes toward the Jews in the Middle Ages: A Casebook* (London: Routledge), pp. 183–209.
Vallentyne, P., and S. Kagan (1997) 'Infinite Value and Finitely Additive Value Theory', *The Journal of Philosophy* 94.1: 5–26.
Waldron, J. (1991) 'Locke: Toleration and the Rationality of Persecution', in J. Horton and S. Mendus (eds), *John Locke: A Letter concerning Toleration in Focus* (London: Routledge), pp. 98–124.

10
Polarized Yet Warranted Christian Belief

David Efird

Does the world and our experience of it constitute evidence for God's existence, or does it constitute evidence against his existence? This question has inspired seemingly endless debate with no rational resolution in sight. To cite some classic contemporary exponents of either side of the debate, on one side, Swinburne (1991a) argues that a variety of aspects of the world and our experience of it constitutes evidence *for* God's existence, an argument he summarizes thusly:

> Why believe that there is a God at all? My answer is that to suppose that there is a God explains why there is a world at all; why there are the scientific laws there are; why animals and then human beings have evolved; why humans have the opportunity to mould their characters and those of their fellow humans for good or ill and to change the environment in which we live; why we have the well-authenticated account of Christ's life, death and resurrection; why throughout the centuries men have had the apparent experience of being in touch with and guided by God; and so much else. In fact, the hypothesis of the existence of God makes sense of the whole of our experience, and it does so better than any other explanation which can be put forward, and that is the grounds for believing it to be true. (1991b)

On the other side of the debate, appealing to other aspects of the world and our experience of it, Rowe (1979) argues that the evil we experience constitutes evidence *against* God's existence. It seems, then, that the world and our experience of it constitutes mixed evidence, that is, evidence which seems to confirm and evidence which seems to disconfirm the existence of God. Now, say that a person has either a fairly confident

belief that God exists or a fairly confident belief that God does not, and she is then presented with the above mixed evidence for God's existence, a situation many students taking philosophy of religion courses find themselves in, as they read Swinburne's and Rowe's now classic arguments. How should the consideration of this mixed evidence affect the confidence level she has in her belief?

It seems that her confidence in her belief should decrease, since, after all, the evidence is mixed. But this isn't what tends to happen. Religious belief, as Elizabeth Weiss Ozorak (1989), Paul Klaczynski and Gayathri Narasimham (1998), and Paul Klaczynski (2000) have shown, is subject to a phenomenon known as 'belief polarization', like many of our other beliefs, such as political beliefs (Chang, 2003), moral beliefs (Epley and Caruso, 2004) and other types of evaluative belief (Lord and Taylor, 2009), where consideration of mixed evidence prompts an increase, rather than a decrease, in subjective confidence in that belief because we assimilate evidence for and against our religious belief in a biased way: evidence which confirms our belief we assimilate in an uncritical way while evidence which disconfirms our belief we assimilate in a critical way. Religious belief, then, tends not only to be preserved in the face of mixed evidence but also strengthened.

Now, it might seem that the polarization of religious belief poses a problem for its warrantedness,[1] since, it seems, religious belief, if polarized, could not then be the product of cognitive processes aimed at the truth. Consequently, religious belief, if polarized, cannot be warranted, or so one might argue. In what follows, after summarizing the phenomenon of belief polarization, I argue that at least one form of religious belief, namely, Christian belief, can be polarized yet warranted, by providing a model on which this is the case.

1. Belief Polarization

Say that you and I disagree about whether capital punishment should be legal. We each have some subjective confidence in our respective beliefs, but not overwhelming certainty. We are presented with 'mixed' evidence, that is, some evidence for making capital punishment legal and some evidence against. How should our confidence in our beliefs change in light of having considered this 'mixed' evidence? It seems that our respective confidences in our beliefs should decrease since the evidence is mixed. However, Charles Lord, Lee Ross and Mark Lepper (1979) found that what seemingly *should* happen to our respective confidences *does not* in fact happen; rather than

decreasing, the subjective confidence we have in our beliefs tends to increase.

Lord, Ross and Lepper discovered this fact in the following way. They measured the strength with which a group of people held their particular belief on capital punishment and then divided the group into smaller groups. They showed the two groups one of two cards, each of which had a statement about the results of a (fictitious) research project written on it:

> Kroner and Phillips (1977) compared murder rates for the year before and the year after adoption of capital punishment in 14 states. In 11 of the 14 states, murder rates were lower after adoption of the death penalty. This research supports the deterrent effect of the death penalty.
>
> Palmer and Crandall (1977) compared murder rates in 10 pairs of neighbouring states with different capital punishment laws. In 8 of the 10 pairs, murder rates were higher in the state with capital punishment. This research opposes the deterrent effect of the death penalty. (Lord, Ross and Lepper, 1979, p. 2100)

The participants were then asked about the strength of their beliefs about the deterrence of the death penalty and the effect the research had on the strength of their belief. Next, the participants were given information on the study described in the card they received, including details of the research, critiques of the research and the researchers' replies to those critiques. The participants were then asked again about the strength of their beliefs about the deterrence of the death penalty and the effect the research had on the strength of their belief. The trial was then re-run on all the participants using the card which supported the opposite belief to that which they had seen initially. Lord, Ross and Lepper found (1) that participants thought that the research which agreed with their original belief was better conducted than the research which disagreed with their original belief and (2) that participants' original belief was held more strongly *both* after reading research which supported their belief and after reading research which conflicted with their belief. Participants' beliefs concerning capital punishment were then polarized having examined mixed evidence regarding those beliefs.

What this study, and many others which confirm its results,[2] shows is that we tend to evaluate evidence for our strongly held views differently

from how we evaluate evidence against those same views. Evidence for our views we tend to assimilate uncritically, without subjecting it to careful scrutiny; evidence against our views we tend to assimilate critically, subjecting it to critical scrutiny, with the result that we take ourselves to have discredited that evidence. The result of this biased processing of evidence is that our confidence level in our views increases when we consider mixed evidence, since evidence for our views, uncritically assimilated, confirms our views, and evidence against our views, critically assimilated and thereby seemingly discredited, further confirms our views. Consequently, the confidence level we have in our beliefs increases; our beliefs become polarized.

2. Polarized Religious Belief

A great variety of our beliefs are subject to this phenomenon, including our religious beliefs (Ozorak, 1989; Klaczynski and Narasimham, 1998; Klaczynski, 2000). Given that our religious beliefs are polarized in this way, what follows regarding their warrantedness? It might be thought that if religious beliefs are polarized, then they cannot be warranted for the following reason:[3]

> Because we assimilate evidence for and against religious belief in a biased way, uncritically assimilating evidence that confirms religious belief but critically assimilating evidence that disconfirms religious belief, which then tends not only to preserve religious belief but also to increase subjective confidence in religious belief, religious belief is not among the proper deliverances of our rational faculties, since it is not the product of properly functioning, truth-aimed cognitive processes whose purpose is to provide us with true beliefs held with an appropriate level of subjective confidence. Rather, religious belief is the result of cognitive dysfunction, of unreliable processes which are not aimed at the truth but rather at the preservation of religious belief, regardless of whether the belief is true or not. Therefore, the polarization of religious belief renders it irrational.

From this objection, it seems to follow that religious belief, if polarized, cannot be warranted since, it seems, the polarization of religious belief renders it irrational.[4] In what follows, I argue that this is not the case; I argue that the polarization of religious belief does not entail its lack of warrant, for at least one kind of religious belief, namely, Christian

belief, by providing a model on which Christian belief is polarized yet warranted.

3. A Model of Polarized Yet Warranted Christian Belief

Following Alvin Plantinga, I take it that

> to give a model of a proposition or state of affairs S is to show *how it could be* that S is true or actual. The model itself will be *another* proposition (or state of affairs), one such that it is clear (1) that it is possible and (2) that if it is *true*, then so is the target proposition. From these two, of course, it follows that the model is possible. (2000, p. 168; emphasis in original)

The model I propose for showing the possibility of polarized warranted Christian belief begins from the model of warranted Christian belief developed by Alvin Plantinga (2000) from the thought of Thomas Aquinas (1948) and John Calvin (1960).

On this model, God has endowed us with a *sensus divinitatis* in order that we may come to know him. The *sensus divinitatis* is a cognitive mechanism which produces belief in God in a wide variety of circumstances, such as the contemplation of the glories of nature, which trigger the disposition to form this belief. On this model, knowledge of God is not arrived at by inference; rather, it is basic, immediate knowledge, akin to perceptual knowledge. This cognitive faculty is aimed at the truth, since God, on this model does in fact exist, and it functions according to its design, that is, in accord with how God designed human beings, in an appropriate environment, the world in which God has created us, which triggers belief in God. As a result, beliefs produced by this faculty are counted as knowledge because they are produced by cognitive processes aimed at the truth, processes which are thereby reliable, and, in fact, it is true that God exists on this model. Because this model is clearly possible, and because it entails that Christian belief is warranted, it shows the possibility of warranted Christian belief.

The model can then be extended to show the possibility of polarized yet warranted Christian belief in the following way. Not only does God desire us to know of his existence, but also he desires to protect that knowledge.

Thus, not only has God designed our minds with the faculty of the *sensus divinitatis* and placed us in an environment which would trigger the operation of that faculty, but he has also designed our minds

and placed us in an environment which would increase our confidence in that knowledge. The environment provides mixed evidence for God's existence, evidence which includes the glories of nature, which tends to confirm belief in God's existence, as well as moral and natural evil, which tends to disconfirm belief in God's existence. But God has designed our minds in such a way that we assimilate this evidence in a biased way such that, rather than the confidence in our belief that God exists decreasing, as one might expect, our confidence increases. So, the environment together with the design of our minds tends to increase our confidence in belief in God, thereby protecting knowledge of his existence.

Therefore, belief polarization, rather than being a product of cognitive dysfunction, is a product of the proper functioning of our cognitive faculties, at least as it functions with respect to belief in the existence of God. Furthermore, the mixed evidence for God's existence provided by the world and our experience of it, rather than constituting a problem for God's existence, as one might think on some versions of the problem of divine hiddenness (see Schellenberg, 1991), this mixed evidence confirms belief in God's existence. Indeed, the evil we experience in the world, because it forms a part of the mixed evidence for God's existence, helps to confirm, rather than disconfirm, belief in God.[5] That God is hidden to some extent, that the evidence for God's existence is mixed, then protects belief in God given the phenomenon of belief polarization. Thus, divine hiddenness together with belief polarization is a way of belief protection for those who do believe in God. Hence, belief polarization is, on this model, a feature of properly functioning cognitive faculties aimed at the truth, functioning according to a design plan, and functioning in an appropriate environment, namely, one which provides mixed evidence for God's existence. Therefore, because this model is clearly possible, and because it entails that polarized Christian belief is warranted, it shows the possibility of polarized, yet warranted, Christian belief. Consequently, the polarization of Christian belief does not entail that it lacks warrant.

4. Objection: The Persistence of Atheism

It might be objected that the model is not possible, since it entails that once we hold atheistic beliefs those beliefs would be protected by the design of our minds, namely, their susceptibility to belief polarization, together with the environment, that is, the mixed evidence for God's existence provided by the world, just as much as theistic beliefs would

be, and this should not be possible given God's desire and ability to have relationship with us.[6] In response, the model can be extended to show this possibility.

The extended model begins with the claim that when we, the human race, that is, were first created, the *sensus divinitatis* was functioning perfectly, and so all of the first humans believed in the existence of God, and that belief was protected by the design of our minds and the environment we were in. However, in addition to endowing us with the *sensus divinitatis*, God also endowed us with free will, free will in the sense of the unfettered ability to determine our own actions. Now, the first humans, Adam and Eve, say, exercised their free will in such a way that they sinned, that is, acted contrary to the will of God. Call this event 'the Fall'. A consequence of the Fall was the corruption of human nature, corruption which included the impairment of the *sensus divinitatis*, such that atheism now became possible. So, given the design of our minds and our environment, atheism as a persistent belief became possible.

Now, this possibility was in no way God's intention for us. But it was not something he could have done anything about, given the fact that he also endowed us with free will, in the sense defined above. So, as a consequence of the Fall, persistent atheism became possible. But, also on this model, at the end of time, God will reveal himself fully, and all will then not only know of his existence but be in union with him, and so atheism will not persist, nor, *a fortiori* will anyone go to a realm of eternal conscious punishment, Hell, as a result of their atheism. So, though persistent atheism is an evil, it is merely a transitory evil and is outweighed by the good of free will. So, it is then possible for God to create us with the *sensus divinitatis* and free will and in an environment that provides mixed evidence for his existence, all of which entails the possibility of persistent atheism. Thus, that the model entails persistent atheism does not show the model to be impossible.

5. Objection: The Unfalsifiability of Polarized Christian Belief

It might be objected that by using a model based on Christian belief in order to defend the possibility of polarized yet warranted Christian belief I have rendered Christian belief unfalsifiable. The authors of the original study which uncovered belief polarization write,

> Our subjects' main inferential shortcoming . . . did not lie in their inclination to process evidence in a biased manner. Willingness to

> interpret new evidence in the light of past knowledge and experience is essential for any organism to make sense of, and respond adaptively to, its environment. Rather, their sin lay in their readiness to use evidence already processed in a biased manner to bolster the very theory or belief that initially 'justified' the processing bias. In so doing, subjects exposed themselves to the familiar risk of making their hypotheses unfalsifiable – a serious risk in a domain where it is clear that at least one party in a dispute holds a false hypothesis – and allowing themselves to be encouraged by patterns of data that they ought to have found troubling. (Lord, Ross and Lepper 1979, p. 2107)

To this objection, I can only reply that just as it must be open to the atheist to be able to assume the possibility of a naturalistic account of the world if the warrantedness of her belief is in question, so must it be open to the theist to be able to assume the possibility of a supernaturalistic account of the world if the warrantedness of her belief is in question. It might then mean that, because of the phenomenon of belief polarization and that the world presents mixed evidence concerning the existence of God, it is impossible to rationally resolve the question of God's existence, which might explain why natural theology and atheology seem to be degenerating research programmes (cf. Lakatos, 1970).

6. Mysterianism about the Existence of God

Those who believe, then, do so by faith, along the lines suggested by John Bishop (2007), who has recently argued that because all publicly available evidence can be interpreted in contrary ways, ways which are either consistent with theism or consistent with atheism, religious believers should be modest fideists, according to which it is sometimes morally acceptable to hold a belief which (1) shapes how a person lives their life, (2) is such that one may genuinely choose to make a practical commitment to its truth, and (3) cannot be evidentially justified within any wider framework of beliefs. Such modest fideism seems to be suggested by the nature of our minds, in particular, how we assimilate evidence for and against our religious views, and the nature of the world, which presents mixed evidence regarding the existence of God.

Colin McGinn (1989) has argued that the solution to the mind–body problem is cognitively closed to us, that is, that it is beyond the capability of the human mind to conceive of it. Similarly, though not exactly

parallel, I suggest that the nature of our minds, the nature of the world, and the nature of the rational resolution of the question of God's existence are beyond the capability of the human mind to decide. Along the lines of McGinn's verdict on whether we can solve the mind–body problem,

> We have been trying for a long time to solve the mind–body problem. It has stubbornly resisted our best efforts. The mystery persists. I think the time has come to admit candidly that we cannot resolve the mystery. But I also think that this very insolubility – or the reason for it – removes the philosophical problem. (1989, p. 349)

the following seems to be true, given the lack of progress in rationally resolving the question of God's existence, a lack of progress explained, at least in part, by the phenomenon of belief polarization and the mixed evidence concerning the existence of God presented by the world:

> We have been trying for a long time to solve the problem of rationally resolving God's existence. It has stubbornly resisted our best efforts. The mystery persists. I think the time has come to admit candidly that we cannot resolve the mystery. But I also think that this very insolubility – or the reason for it – removes the philosophical problem.

If this is correct, then a 'mysterian' (a term coined by Owen Flanagan in his 1991, p. 313) position on the existence of God seems true: that there could (epistemic sense of 'could') not be a rationally compelling argument, one that should convince everyone concerned, for theism or for atheism, because of the nature of our minds and the nature of the world. There is, then, no *philosophical* problem of God's existence.[7]

Notes

1. By 'warrant', I mean whatever it is that distinguishes true belief from knowledge.
2. For further studies, see Gilovich, 1991, ch. 3 and Lord and Taylor, 2009.
3. This style of objection, that *religious* belief is the product of cognitive processes not aimed at the truth, is familiar from Sigmund Freud (1961) and Karl Marx (1964).
4. In a recent article, Thomas Kelly (2008) considers the epistemological implications of belief polarization, and he argues that the polarization of belief does not affect the reasonableness of holding that belief, for those who are unaware of its being polarized. He writes, 'I think that if one's beliefs are ones that it

would otherwise be reasonable to hold in light of one's total evidence, then the fact that it is a highly contingent, fragile matter that one has this particular body of total evidence rather than some other is not enough to undermine the reasonableness of believing as one does' (2008, p. 629). However, Kelly goes on to argue that the polarization of belief does affect the reasonableness of holding that belief, for those who are aware of its being polarized. He writes, 'In general, the fact that distorting or biasing factors played a role in one's arriving at total evidence *E* does not make it unreasonable to believe in accordance with *E*, provided that one is unaware of the operation of those factors; what would be unreasonable would be to fail to adjust one's views upon learning of the role played by those distorting or biasing factors' (2008, p. 629). A consequence of Kelly's conclusion is that, for beliefs we believe to have been polarized by the biased assimilation of evidence, we should be less confident in those beliefs; indeed, it seems we should tend towards agnosticism regarding the subject matter of those beliefs. Now it has been shown that a great variety of our beliefs, particularly our evaluative beliefs, such as political and moral beliefs, are polarized in this way (see Lord and Taylor, 2009). Therefore, if Kelly is right, we, that is, you and I, should then tend towards agnosticism on political and moral matters, amongst many other such matters. This is a striking consequence, and, if correct, would transform political and moral debate. Furthermore, not only are political and moral beliefs subject to belief polarization, but also religious beliefs as well (see Ozorak, 1989; Klaczynski and Narasimham, 1998; Klaczynski, 2000). Therefore, if Kelly is correct, then, because religious belief is polarized, those readers who have religious beliefs should now be less confident of those beliefs, even becoming religious agnostics.

5. This account of evil provides a theodicy similar to that of John Hick's (1966) soul-making theodicy, a sort of confidence-making theodicy, which justifies God's placing humans in an environment containing evil for benefiting the relationship between God and humans.
6. I am grateful to Yujin Nagasawa for raising this objection.
7. I am grateful for Yujin Nagasawa's most helpful comments on a previous draft of this essay.

References

Aquinas, Thomas (1948) *Summa Theologica*, vol. 1, trans. Fathers of the English Dominican Province (New York: Benziger Bros.).

Bishop, John (2007) *Believing by Faith* (Oxford: Clarendon Press).

Calvin, John (1960) *Institutes of the Christian Religion*, trans. Ford Lewis Battles and ed. John T. McNeill (Philadelphia: Westminster Press).

Chang, Chingching (2003) 'Party Bias in Political-Advertising Processing: Results from an Experiment Involving the 1998 Taipei Mayoral Election', *Journal of Advertising* 32: 83–92.

Epley, Eugene, and Nicholas Caruso (2004) 'Egocentric Ethics', *Social Justice Research* 17: 171–87.

Flanagan, Owen (1991) *The Science of the Mind* (Cambridge, MA: MIT Press).

Freud, Sigmund (1961) *The Future of an Illusion*, trans. and ed. James Strachey (New York: W. W. Norton).

Gilovich, Thomas (1991) *How We Know What Isn't So* (New York: The Free Press).
Hick, John (1966) *Evil and the God of Love* (New York: Harper and Row).
Kelly, Thomas (2008) 'Disagreement, Dogmatism, and Belief Polarization', *The Journal of Philosophy* 105: 611–33.
Klaczynski, Paul A. (2000) 'Motivated Scientific Reasoning Biases, Epistemological Beliefs, and Theory Polarization: A Two-Process Approach to Adolescent Cognition', *Child Development* 71: 1347–66.
Klaczynski, Paul A., and Gayathri Narasimham (1998) 'Development of Scientific Reasoning Biases: Cognitive Versus Ego-Protective Explanations', *Developmental Psychology* 34: 175–87.
Lakatos, Imre (1970) 'Falsification and the Methodology of Research Programmes', in Imre Lakatos and Alan Musgrave (eds), *Criticism and the Growth of Knowledge* (Cambridge: Cambridge University Press), pp. 91–196.
Lord, Charles G., Lee Ross, and Mark R. Lepper (1979) 'Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence', *Journal of Personality and Social Psychology* 37: 2098–109.
Lord, Charles G., and Cheryl A. Taylor (2009) 'Biased Assimilation: Effects of Assumptions and Expectations on the Interpretation of New Evidence', *Social and Personality Psychology Compass* 5: 827–41.
Marx, Karl (1964) 'Contribution to the Critique of Hegel's Philosophy of Right, Introduction', in idem and Friedrich Engels, *On Religion*, trans. Reinhold Niebuhr (Chico, CA: Scholar's Press).
McGinn, Colin (1989) 'Can We Solve the Mind–Body Problem?', *Mind* 98: 349–66.
Ozorak, Elizabeth Weiss (1989) 'Social and Cognitive Influences on the Development of Religious Beliefs and Commitment in Adolescence', *Journal for the Scientific Study of Religion* 28: 448–63.
Plantinga, Alvin (2000) *Warranted Christian Belief* (New York and Oxford: Oxford University Press).
Rowe, William L. (1979) 'The Problem of Evil and Some Varieties of Atheism', *American Philosophical Quarterly* 16: 335–41.
Schellenberg, J. L. (1991) *Divine Hiddenness and Human Reason* (Ithaca, NY: Cornell University Press).
Swinburne, Richard (1991a) *The Existence of God*, rev. edn (Oxford: Clarendon Press).
Swinbrune, Richard (1991b) 'The Justification of Theism', *Truth Journal*; available online at www.leaderu.com/truth/3truth09.html (accessed 4 July 2011).

Part VI

The Compatibility of Science and Religion

This page intentionally left blank

11
Freedom, Science and Religion

Katherin A. Rogers

1. Introduction: Christian Freedom

The interrelated questions of free will, moral responsibility, self-reflectiveness and formation of character are central for the Christian life. In recent decades experimental psychologists, and philosophers working on the data they provide (henceforth shortened to just 'experimental psychologists'), have produced exciting evidence related to these issues – evidence about how we go about choosing, thinking about our choices and developing our characters. This evidence and future work along these lines can prove invaluable in helping those who are interested in choosing wisely and well and in building good character. For example, recent experiments show that it is surprisingly common for us to misremember or to fail to grasp just what it is we have done or are doing and why. So experimental psychology has much to offer to the Christian believer. However, some experimental psychologists evince an attitude bordering on contempt towards their fellow human beings and an eagerness to go beyond what the evidence warrants to make sweeping claims about human nature including that we do not have free will or robust characters. These claims are antithetical to Christianity, or indeed to any view which attributes to human beings a special dignity as free, rational agents. I will not offer a lengthy rehearsal of the arguments reconciling the experimental evidence with the existence of free will and character (Alfred Mele [2009] has recently produced a good book on the topic). Rather, I will sound a cautionary note concerning experimental psychology. I will defend the claim that there is a practical danger in denying human freedom by appealing to evidence which can be found in the way experimental psychologists go about their business. This is not a condemnation of the experimental

approach. I will suggest ways in which future experiments can be conducted and evaluated which may help us understand and perhaps even improve upon our situation as intentional actors, but which avoid some of the worrisome aspects apparent in present work.

The assumption of human freedom is central to the Christian conception of the relationship of the human being to God. Doctrines concerning sin and salvation, reward and punishment, depend upon the underlying claim that humans are free and morally responsible. More fundamentally, it is in being free, rational agents, with some control over the sort of people that we are, that we are said to be exceptional among animals as specially made in the image of God. This insistence on human dignity entails two practical corollaries concerning human equality. The Christian claim, most counter-intuitive to the classical Greco-Roman culture within which Christianity first grew, is that all human beings – even the poor or enslaved! even women and children! – are at the core metaphysically equal. We are each of tremendous value having been made in the image of God (Rogers, 2008, pp. 56–60). The practical consequence is that, on a fundamental level, everyone deserves equal respect.

The Christian worldview entails a practical human equality through another important route. If we are indeed such metaphysically elevated creatures, how is it that we always and everywhere fall short of behaving as well as we could? The story of humanity contains much that is dismal and vile. The Christian explanation is that, at the dawn of human history, things went badly wrong. Humans turned against God and introduced a state of war between humans and God, between one human being and another, and even within the human being itself. The war goes on in all of us under the name of Original Sin. Sometimes one reads that contemporary anthropology or sociology or psychology has disproved this doctrine. But alternate explanations for the ubiquity of evil, for example that we behave badly due to our evolutionary history, do not really capture the phenomenology of moral evil (Rogers, 2002, p. 79). If we allow, as the doctrine of Original Sin has it, that we are none of us as good as we ought to be, we should recognize a basic equality among human beings. We should see that we are flawed and so no one among us should set himself over the rest as somehow ultimately and intrinsically superior.[1] If the traditional doctrine of human freedom is explicitly and widely abandoned, these corollaries concerning equality might go with it.

Happily, the sweeping claims regarding the lack of freedom and character go far beyond what the evidence actually warrants. Often, the

experimental psychologist does not even offer a definition or description of free will. On almost any philosophically plausible understanding of free will the experiments do not really speak to the question of whether or not such a thing constitutes part of the human being's mental landscape. As to character, there are myriad ways of understanding the evidence which purports to show that character has little to do with behaviour, such that a more thorough examination of the range of possibilities might well preserve our traditional understanding of character. Thus the evidence does not actually threaten the Christian view of human nature. As so often happens, the claim that 'science conflicts with religion' is unfounded. Still, the overly ambitious claims of some experimental psychologists may pose a danger; a danger which, I will argue, is evidenced in their work.

2. Wegner and Doris: Evidence against Freedom and Character

I take as my prime examples of cause for concern Daniel Wegner's provocative *The Illusion of Conscious Will* and, to a somewhat lesser extent, John Doris's *Lack of Character.* Each proposes, at book length, a sweeping revision of our understanding of human nature. My discussion of their views will be sketchy, I do not say that all work in experimental psychology exhibits the flaws I point to in Wegner and Doris, and by 'Doris' and 'Wegner' here I do not mean the men, but rather the voices in the books. My criticism is aimed at attitudes and assumptions which are implied by the text, but it may be that Doris and Wegner, the men, do not now hold the views which I associate with their names. It is possible that they never did. I will attribute unseemly attitudes to Doris and Wegner – the voices in the text – but one of the themes I develop is the importance of charity and respect for one's object of study, and I do not want to violate my own principle. I take it, though, that the only respect we owe the text is a fair interpretation.

The thesis of *The Illusion of Conscious Will* is that human beings do not have the sort of free will which we usually take ourselves to have. Wegner does not spell out what he means by free will and his positive view of what is going on in the human being is unclear and seems to shift over the course of the book. But one major theme is that our 'conscious will' – again not carefully explained, but including our conscious motives and intentions and choices – is merely epiphenomenal. There are body events (Wegner seems to be a mind/body dualist) which produce conscious events including intentions and so forth and which also

produce subsequent overt bodily behaviours. Since the conscious events are usually closely followed by these subsequent behaviours, we suppose that the behaviours are intentional actions produced by our conscious will. But in fact the apparent actions and the preceding conscious 'willings' are merely constantly conjoined, they are not causally connected. But if all that mental activity is epiphenomenal then we are not free as we thought we were (Wegner, 2002, p. 68). This is an extremely radical position. Should you, right now, decide to raise your right arm (the reader may try this on his own) and then up goes your arm, it is not that the decision is the, or even a, cause of the raising of the arm. Some earlier body event of which you are unaware is the cause of both. Were Wegner's analysis correct, it would rule out not just libertarian theories of freedom, but most, if not all, compatiblist theories. Most standard compatibilist views allow that agents may be free and responsible, even if their choices are causally determined, as long as their actions are voluntary, the consequence of their own deliberations, interests and desires. Wegner's theory denies that agents' deliberations and so forth play any causal role in the production of their actions.

What is Wegner's evidence for this extreme theory? Wegner leans heavily on experiments by Libet which, apparently, show that a special sort of brain activity precedes the choice (or urge) to flex a finger. Does this prove that the unconscious brain event, and *not* the choice, constitutes the sufficient cause of the finger flex? No, the choice might be part of the chain of causes of which an earlier link was the unconscious brain event (Mele, 2009, p. 11). Libet himself holds that freedom may be evidenced by the fact that apparently between the brain event and the flex the agent might 'veto' the choice and not flex. And even a robust libertarian brand of freedom, such as that proposed by Robert Kane, would *expect* to find special brain activity before a choice. On Kane's analysis, a choice follows upon a 'struggle' to pursue one of two incompatible desires or intentions (Kane, 1996, pp. 112–15). The brain event in Libet's experiments may correspond to this struggle.

The bulk of Wegner's evidence consists in experiments and observations which demonstrate a disconnect between the actual bodily behaviours of an agent and what the agent believes about what he or she will do or has done. This is supposed to show that it is not the agent's conscious will that produces bodily behaviour. Many of the experiments and observations are extremely interesting and informative, but they do tend to focus on the odd, the anomalous and the truly bizarre. That there is a disconnect between intention and action in the unusual case does not show that there is *always* a disconnect. Rather the opposite;

what makes a case *un*usual is that the connection which seems to be there in the usual case is missing. Wegner concludes his argument with a review of Daniel Dennett's criticisms of libertarian freedom, which criticisms do not lack for responses in the literature and which, in any case, do not support Wegner's epiphenomenalism. The case for his radical position is excessively thin.

The same can be said for Doris's argument in favour of situationism, the thesis that character as commonly understood rarely if ever exists and that behaviour is best explained by elements in the situation immediately surrounding an agent's actions. Ordinarily we suppose that people have character traits, strong tendencies to act in certain ways, including virtues and vices. We take it that character plays a significant role in producing and explaining behaviour. In fact, says Doris, minor factors in the agent's situation play a much larger role than whatever character traits we might attribute to a person. Character has so little explanatory value that it is reasonable to conclude that character as traditionally understood exists rarely, if ever. Doris chooses the virtue of compassion as his test character trait and reviews a series of experiments and observations wherein, he holds, the compassionate act, or the failure to act compassionately, is most often elicited by a small factor in the situation. But if something minor can make people act more or less compassionately, then probably the subjects do not have a strong tendency towards behaving in a compassionate way. Few people, then, can be said to have compassion as a virtue, and it is the same, presumably, for other character traits.

As with Wegner's evidence, the experiments and observations are extremely interesting and informative. But, again, the conclusion far outstrips the evidence. One problem, on which I will elaborate below, is that one person's minor factor may be another's powerful motivator. Moreover, the choice of compassion may be a poor one for making the broader situationist case. Even if very few people have a fixed tendency towards compassionate activity that would not demonstrate that other character traits are not present robustly in great swathes of the population. For example, raising children is a long and difficult task, but most parents do not let small things interfere with their exercising parental solicitude. Some do, of course, so we cannot say it is a necessity of human nature to be solicitous towards one's children. Might it not fit the definition of a virtue? What about gluttony? The gluttons of my acquaintance are rarely deterred by anything, large or small, from eating too much and devoting too much time and energy to thinking about, watching television shows about and pursuing food. It seems

wildly counter to ordinary experience to hold that no one or only very few people have gluttonous characters.

3. A Dangerous View

Wegner and Doris are overly eager to conclude that human agents do not have freedom and character. And isn't there a danger of dire consequences in advertising these conclusions? If so, they should be disseminated only if there is powerful evidence in their favour. Dennett, who is often quoted on this point by experimental psychologists, including Wegner, assures us that the danger is the invention of an overwrought imagination. Although we do not have the sort of freedom that would ground 'before-the-eyes-of-God' guilt, there is still reason to hold people responsible and punish them. We do so in order to prevent and deter further crime. So even in the absence of a robust freedom, such as libertarians insist upon, we have as much freedom as we should want or need to maintain the fabric of society (Dennett, 1984, pp. 156–72). Even the radical claims of Doris and Wegner are consistent with the thesis that punishment may change behaviour (though Wegner will have to tell an odd story absent any effect of conscious will on body behaviour). So both could hold, along Dennett's lines, that their claims do not undermine ascriptions of freedom and responsibility in any harmful way.

But, from the perspective of the Christian, Dennett rather misses the point. True, much of the free will literature revolves around justified punishment. But the question of punishment is a by-product of the more fundamental question of what sort of things we are. In the Christian tradition the *point* of having free will is that we should be able to be good on our own (in some very limited, creaturely way) and so be the best images of the divine that we can be. If we are not free enough for before-the-eyes-of-God guilt, then we are not free enough for before-the-eyes-of-God praiseworthiness. We are not the exceptional, made-in-the-image-of-God creature that the Christian tradition has taken us to be. We do not have the human dignity and the elevated value which the religious perspective ascribed to us. Dennett's assumption that the whole reason for holding people responsible and punishing them is so that we may use them to benefit the rest of us illustrates the point nicely.

This denial of human freedom and character carries with it the danger of rejecting the corollaries about equality – that all human beings need to be treated with respect and that no one should set himself up as

intrinsically better than the mass of men. This danger becomes sharply focused when the denial of freedom comes from those who study the causes of human behaviour. History demonstrates that people are always looking for ways to control others and that the 'others' are usually the worse for it. Knowledge is power. A standard impetus for the study of the natural world has been to exert control over it. Could the study of human beings through experimental psychology inspire and enable those who would like to do the controlling? Not necessarily. If the experiments are done respectfully and the conclusions are drawn cautiously, experimental psychology need not encourage the power seekers. On the contrary, in helping all of us understand the connections between our desires, motivations, intentions, actions and characters, experimental psychology may confer upon the individual agent more *self*-control and autonomy than he could have had otherwise. But what if the respect and caution are lacking? Suppose the extreme claims of Doris and Wegner were to become the consensus in our society? Suppose we come to accept that no one is ever free and responsible except in Dennett's punishment-serves-a-social-purpose way? Then, in attempting to control others through behavioural engineering, we do not impose some new restraints on them; we simply replace one set of stimulae thrown up by a blind nature with a new set produced and monitored by the expert in human behaviour. And isn't that an improvement? That, at least, is how the argument of the power-seekers might proceed.

C. S. Lewis in *That Hideous Strength*, published in 1943, has one of his villains explain the point of studying human beings: 'If Science is really given a free hand it can now take over the human race and re-condition it: make man a really efficient animal' (1997, p. 383). Later in the book an even worse villain amplifies: 'You know as well as I do that Man's power over Nature means the power of some men over other men with Nature as the instrument' (1997, p. 519). B. F. Skinner in *Walden Two*, published five years later, puts almost the same words in the mouth of his *hero*. Frazier, Skinners's mouthpiece in the book, defends a contemporary version of Plato's old, anti-democratic claim that the best society is the one in which the expert is in charge: 'Are the people skilled governors? No. And they become less and less skilled, relatively speaking, as the science of government advances . . . when we've once acquired a behavioral technology, we can't leave the control of behavior to the unskilled' (1976, pp. 250–1). Later (and one must give Skinner credit for honesty) Frazier notes that there is a 'curious similarity' between him and God, and then – as if he had just been reading *That Hideous Strength* – says, 'When we ask what Man can make of Man, we don't

mean the same thing by "Man" in both instances. We mean to ask what a few men can make of mankind. And that's the all-absorbing question of the twentieth century. What kind of world can we build – those of us who understand the science of behavior?' (p. 279). Skinner's extreme behaviourism is now passé, and one might suppose that concern about experimental psychology is paranoia induced by exposure to *A Clockwork Orange* at too impressionable an age. I will argue, to the contrary, that there is evidence in Doris and Wegner that fear is warranted.

4. 'Using' the Subject

My evidence is of three sorts. First, there is a problematic willingness to 'use' the participant in the experiment in various ways. Second, there is a lack of charity towards the participant in interpreting the experiment. Both of these points speak to an absence of the sort of respect that the traditional view of the free human person called for. Third, and most importantly, Doris and Wegner set themselves above the mass of humankind in that their operating assumption is that they are not subject to the sort of limitations they have discovered in all the rest of us. In discussing each sort of evidence, I will point to ways in which experimental psychology might proceed without the worrisome factor.

Some experimental psychology involves setting up situations and then observing how passers-by respond. Some involves experiments in which the participants know that they are part of a study. In neither case are people coerced to participate. The Institutional Research Boards (IRBs) at universities are charged with seeing that experiments using human beings evince a 'respect for persons' which ensure that participants are not coerced into participating. At my university a requirement in many psychology courses is participation in studies, but substitution of a written paper is usually allowed, and, in any case, no one is forced to take psychology courses. Coercion to participate is not the problem. But even a voluntary participant may be 'used' in ways that are troubling. The famous Milgram obedience experiments cited by Doris and Wegner illustrate two areas of concern. Apparently the contemporary community of experimental psychologists disapproves of these experiments as they were originally done, but they continue to conduct experiments which demonstrate the objectionable features to which I will point.

In the Milgram experiments paid volunteers were told that they were participating in an experiment concerned with the effect of punishment on memory. A lab-coated 'experimenter' explains to the volunteer that he is to administer a shock to a fellow volunteer if he gives

incorrect answers to questions and that, as the number of wrong answers increases, the intensity of the shocks increases. Actually the other 'volunteer' is a confederate who does not feel any shock. As (supposedly) more shocks of more intensity are administered the confederate begins to scream, to claim that his heart is bothering him, and to beg to be released. If the 'shocks' continue past a certain point the confederate falls silent. No response is considered a wrong answer and the volunteer is told to continue the shocks. If, at any point in the proceedings, the volunteer balks at administering another shock, the 'experimenter', in a firm but polite tone, asks the volunteer to continue, says the experiment 'requires' that he continue and so forth. At any time the volunteer can refuse to administer more shocks with no reprisal. The dismaying result was that the majority of volunteers were willing to administer shocks past the point at which the confederate had fallen silent (Doris, 2002, pp. 39–42).

An IRB today would probably not permit a recreation of the whole Milgram experiment, though a recreation of part of the experiment was recently conducted by Jerry M. Burger. It produced roughly similar results (Burger, 2009). Burger noticed, in reviewing the Milgram data, that a large percentage of those who were willing to administer shocks up to 150 volts were willing to go all the way. He recreated the experiment, with the difference that once a participant agreed to go past the 150-volt threshold, the experiment was stopped so that the participant did not have to go through the ordeal of deciding whether or not to administer more severe shocks. Also, the participant was immediately debriefed. The IRB would likely judge, of the original experiment, that the harms might outweigh the benefits – the harms being the immediate stress and anxiety suffered by the participants, any subsequent psychological harm to them, and the possibility of harm to their reputations if their behaviour during the experiments should become public knowledge. These are well-taken concerns, but they do not exhaust the worrisome features. A further cause of concern in this experiment is that the subject is simply being lied to. Many, if not most, of the experiments recorded by Doris and Wegner involve lying to or otherwise deceiving the subjects, as does the recent recreation of the Milgram experiment. It is a curious phenomenon – at least to the philosopher steeped in the Christian tradition – that experimental psychologists seem to take it for granted that lying to people is acceptable in the setting of the experiment. This is not 'play-acting' lying, such as one might do on stage, since the participants need to be genuinely deceived in order for the experiment to provide the sort of evidence being sought. But deceiving

someone is a classic instance of 'using' them. It is treating one as a means to the ends of the deceiver. Doesn't the liar express an unwholesome attitude of superiority over the participant? A common principle upon which IRBs operate today is that deception is permissible (assuming other principles of beneficence and justice are met) if the participant is immediately debriefed after the experiment. This does seem to mitigate the concern, but does not allay it entirely. Presumably none of us like being lied to by plumbers or waiters or policemen. (Many of the untruths in the mouths of politicians probably fall into the 'play-acting' category, since almost no one is really deceived.) And even if the waiter tells us later that he lied, it is not clear that that undoes the wrong. (I use this example because a student of mine argued that deception in experimenting is permissible because it is no more harmful than the lying he did as a waiter, always recommending the most expensive item so he would get a big tip.) Is experimental psychology not governed by basic rules of honesty which one would hope to apply in other professions?

One might respond that lying to the participant in the experiment is permissible as long as it does not do them any harm – especially if they are immediately debriefed. A first response is that lying to someone, even if it does them no harm, is still wrong. Many well-developed and defensible ethical theories hold as much. Experimental psychologists tend (at least this is the impression one gets from the literature and discussion of the IRB principles) to be consequentialists. And this should raise questions in that most versions of consequentialism, as they address the human condition, assign the locus of ultimate value to something other than the intrinsic worth and dignity of each human being. A second point to note is that, even if a deceiving experiment did not harm the participant, it might harm the experimenter. What if Aristotle is right and Doris is wrong? What if we do have characters and we formed them at least in part through our actions? Burger's understanding of the Milgram experiments is of interest on this point. He takes the evidence to show that someone who continued to give shocks up to the 150-volt level was more likely to continue to the bitter end. This might be because someone who is willing to give a shock of 150 volts is already the sort of person who would be equally willing to give a much more damaging shock. But an Aristotelian could take the evidence to indicate that choosing to go to 150 volts *made one* the sort of person willing to go much further. If so, minor choices can have significant consequences for the agent themselves. Surely it is not a good thing to be a liar. One who lies often as part of the job, even if the lying is innocuous in the workplace context, may become a more general liar. Thus one could

argue that the experimenter is put in danger. The IRB, I am told, does not oversee the well-being of the experimenter.

If the above point is plausible, then the deceived participant in an experiment may be harmed in ways which IRBs do not anticipate. The IRB takes note of stress and anguish which might be suffered during the experiment, and it looks to the issues of risks of psychological harm or reputational harm. From the Christian perspective there is more at stake, however. If we are interested in studying freedom and character, the most obvious way to construct an experiment is to place a participant in a situation involving significant moral choice. That is what the original Milgram experiments did, and, given Burger's thesis that the 150-volt point is the decisive moment for whether or not a participant is likely to proceed the whole way, the recreated experiment includes an almost equally significant moment of choice. Many of the other experiments which Wegner and Doris cite place participants in morally significant situations. Let us grant that actually shocking someone badly, possibly to death, in an experiment concerning the effects of punishment on memory would be evil. (I do not know that 'evil' is a word much used in the psychology literature, but, at least in the USA, experimentation using human subjects is to be guided by the report, issued in 1979, of the Belmont Commission. The Commission was charged with establishing basic *ethical* principles. I assume that means that the scientific community is to be guided by what is right and wrong, good and evil. Evil, then, should be allowed as a reality in the universe of experimental psychology.) From the perspective of the moral choices and character of an individual, if the participant believes the memory experiment to be real, then, even in the fake version, the subject is participating in evil. The Christian standardly holds that doing wrong and making bad choices does great harm to the agent. One ultimate hope of the Christian is to be praiseworthy-in-the-eyes-of-God. But you become a murderer by murdering. You become a torturer by torturing (Rogers, 2007). The subjects in the Milgram and Burger experiments were put at risk of harm – they were placed in moral peril, danger of becoming wicked people – just by being thrust into a situation in which they were sorely tempted to do the wrong thing.

True, the ordinary vicissitudes of life put us in positions where we have to make moral choices. But the fact that some event is sure to befall us as we go about our business does not entail that it is permissible for our fellows to engineer that event. Think of death. And, yes, the few who refused to continue with the shocks may have ultimately benefited by doing the right thing. But the results of the experiment

indicate that overall the subjects incurred more harm than good, and, in any case Milgram and Burger could not gauge how much harm would be done. We might defend the experiments by noting that the harm of making the bad choice was up to the subject themselves. True, but they could not have made the choice had they not been put into the situation. If the vicissitudes of life put you in situations where you are sorely tempted to do the wrong thing, that is bad luck for you. That does not give the experimenter carte blanche to construct some bad luck for you. The Christian must grant that God may, from time to time, for good reasons of His own, say 'No' to our prayer that He 'lead us not into temptation'. But God is a special case and to argue that the experimenter may do what God may do would prove my point about the dangerous attitude associated with denying freedom and character. Perhaps the most worrisome point about the risk of moral harm to the tempted subject is that it does not seem even to occur to experimental psychologists and IRBs. Certainly Wegner and Doris do not mention it.

If lying and putting people in morally tempting situations may cause harm, the experimental psychologist could respond that the harm is outweighed by the good of the knowledge produced. The non-consequentialist answers that it is still wrong to use people. But even the consequentialist must grant that these are the sorts of harms and goods it is exceedingly difficult to balance, especially if we add the harm done to the agent who chooses evil. The 150-volters in Burger's experiment may have done themselves permanent and serious damage without the IRB even considering the possible risk. Must experimental psychologists simply stop studying human freedom and character? It will be difficult to construct fruitful experiments which put people in morally interesting situations, but which do not deceive them. Knowing they are being tested morally is almost certain to limit the value of the participants' responses. But there are other avenues which the experimental psychologist might pursue. I will mention three, but I presume there are many more.

Let us take the question which motivated Milgram and Burger. In Nazi Germany many otherwise apparently ordinary folk engaged in hideous behaviour, often in obedience to orders they might have disobeyed without terrible consequences. Could ordinary Americans in the 1960s or the 2000s do the same? Can we study that question without using deception and placing participants in moral peril? First, it might be possible to construct experiments which do not need deception. They might involve choices which do not have moral significance, but which are somewhat similar, and so might shed light on how various factors

affect moral decisions. One might ask agents to make aesthetic choices or to make judgments about what fictional characters ought to do. If it is important to include the fellow with the lab coat, but without deception, one group of subjects could meet an expert who will tell them how he or she thinks they ought to choose, and then the control group could be left to its own devices. With careful analysis of the relationship between the aesthetic choice or the fictionalized choice and an actual moral choice this could provide at least some evidence.

Second, there are first-person reports. We might compile lists of folks, past and present, who have done hideous things upon being instructed to do so. We could find out about them from written records – court transcripts might be an example – and testimony of family and friends. Do they seem relatively ordinary? Then we could amass first-person reports of what they had done and why. If they are alive we could interview them. If not, perhaps they have left testimony and letters.

Third, once the psychologist's question has been carefully framed, simply observing people as they go about their business or *have* gone about it in the past might provide good evidence. This seems obvious in the Milgram case. A glance at the French Revolution, Russia under Stalin, or China under Mao makes it abundantly clear that hideous behaviour on the part of average people in obedience to orders – especially if a patina of Progressivism and Science is added – is not unique to Nazi Germany. The evidence of history is overwhelming and one who insists that *it couldn't happen here* is naïve or blinded by jingoism. Milgram really didn't need to do his experiments in order to answer his question. There may be other questions which, if suitably framed to focus on relevant evidence, could be answered from observation without the need to use deceptive experiments.

5. Interpreting Evidence Uncharitably

A second sort of evidence that there may be something worrisome in the attitudes of experimental psychologists towards their subjects lies in their frequently assigning an uncharitable interpretation to the results of an experiment when a charitable interpretation is available. I offer two examples, but there are many more. Doris, to defend situationism and undermine the position that people have robust characters, needs to show that behaviour is significantly affected by relatively minor variations in situation. This is because we would expect extreme situational factors to override even robust character traits. For example, even the most compassionate person is likely to behave without compassion

under threat of death. Doris notes that in the Milgram case there is no threat of death, punishment, or indeed negative consequences of any sort, made against the participant should he stop the shocks. Doris then takes the Milgram experiments to show that a minor factor, the instruction of the 'experimenter', has a profound effect on the behaviour of the subjects. So these experiments provide evidence of lack of compassion. But there is a more charitable interpretation (which is still consistent with my point above concerning recent history's unhappy tale about human capabilities): in our society science is taken as the paradigm of proper knowledge acquisition. The scientist is treated with enormous respect. In some contexts – doubt about evolution, for example – those who do not accept the pronouncements of science are held up to ridicule. And given the wonderful advances science has made in the last five hundred years this attitude of respect and deference towards the scientist and their discipline seems appropriate in many or most cases. Add that fear of embarrassment is an extremely powerful motivator. We very often remember our embarrassing moments with much more vivid pain and regret than we suffer when we remember those moments in which we failed morally. Perhaps the subjects in the Milgram experiments, having deeply internalized our contemporary attitude towards science with the entailment that one who is sceptical about the scientist and their project is deserving of contempt, found it extremely difficult to disobey. It is interesting that Milgram focuses on authority, and not the authority of the scientist specifically. Milgram did conduct his experiments in various venues to test the effect of the impact of the institution, but the 'scientist' experimenter was always present (though sometimes through the telephone). What if the experimenter had feigned a stomach disorder and had handed his clipboard – with a statement about delegating authority – to a confederate disguised as a passing janitor? It is at least possible that the subjects in the Milgram experiments were, by and large, compassionate people suffering under severe social pressure not to doubt or disrespect the scientist. If so, their lack of compassion was not the consequence of a *minor* situational factor.

Another example of lack of charity comes from Wegner. The subject is told he is taking a test involving speed typing. He is at a computer keyboard in a room where others (confederates) are at work as well. The subject types until he is told to stop. He is then accused of damaging the computer by pressing the wrong key. In fact he did not, and usually subjects initially deny doing the deed. Then a confederate claims he saw the subject hit the wrong key. 'Those whose "crime" was ostensibly witnessed became more likely to sign a confession. . . . We are not infallible

sources of knowledge about our own actions and can be duped into thinking we did things when events conspire to make us feel responsible' (Wegner, 2002, pp. 10–11). The claim that this result, or similar results from a million such experiments, would provide evidence that consciousness is epiphenomenal is bizarre. But what worries me here is Wegner's view that the subject is 'duped'. 'Dupe' in the dictionary is synonymous with 'fool'. True, the subject is mistaken. But given the evidence available to him, the subject has in fact come to what looks to be the correct conclusion. Is it more probable, epistemically, that you might hit the wrong key when typing quickly, or that both the person conducting the typing experiment and a total stranger would conspire against you to accuse you of doing something you didn't do? A reasonable person should probably *believe* that he had damaged the computer and perhaps even try to remember doing so. Someone who stuck to the story that he *hadn't* damaged the computer would likely be lacking in appropriate humility, motivated by fear of responsibility, and/or not very committed to telling the truth as he sees it. Wegner offers many similar experiments throughout his book and never distinguishes between making a mistake due to some sort of cognitive failure (his example of 'table-turning' where people believed that spirits were making a table move when it was the people themselves doing it might fall into this category) and making a mistake when the available evidence renders the mistaken belief epistemically justified. Lumping everyone together as equally 'duped' fails to respect the abilities of those whose cognitive processes are functioning properly, but have come up against the weird and anomalous situation. And since this failure is pervasive in Wegner's work it suggests a worrisome contempt for the subjects. My prescription, then, is that in interpreting the results of experiments, ascribing negative characteristics such as gullibility or a lack of compassion is appropriate only after more charitable interpretations have been tested and shown to be less plausible.

6. Assuming Superiority

My final sort of evidence is the most serious. Doris and Wegner set themselves outside of and above the rest of humankind in drawing their conclusions. This is more obvious and striking in Wegner's case than in Doris's, so let us look at the latter quickly and then devote more time to the former. The case against Doris requires an assumption which seems intuitively plausible; the proper practice of science (and philosophy – Doris is actually a philosopher assessing the scientific evidence) requires

certain virtues such as perseverance, courage, honesty, humility in the face of the evidence and perhaps many more. Doris's theory, situationism, holds that human beings do not have virtues and vices, but behave as they do because of minor variations in situation. Were that the case we might expect the scientist whose research has turned up a conclusion that poses a problem for him to be honest if he's found a five-dollar bill (there is an old experiment supposedly showing that people acted more compassionately when they had found a dime) but not honest if he's stepped on chewing gum. The successful practice of science would be rendered impossible. Doris is engaging in professional philosophy and his trust that the experiments he reports were conducted properly suggests that he takes himself and the scientists he cites to be successful practitioners of science and philosophy. But then Doris and his confreres must possess the virtues necessary for the job. Thus Doris seems to exempt himself in concluding that people do not have characters.

The case with Wegner is more obvious and more extreme. Regarding Doris one *could* argue that the 'scientific' virtues are not truly required for the successful practice of science or that the virtues in question are not really robust character traits. So, *possibly* Doris could consistently apply his conclusions to himself. The same cannot be said for Wegner. Wegner holds that conscious will – that is intentions, motivations and so forth – is epiphenomenal. We are deluded in thinking that our intentions even partially cause our overt actions. Our actions, and possibly our intentions themselves, are caused by preceding body events of which we are unaware. If this thesis were correct, then no one could ever conduct experiments. A representative dictionary (Webster's) defines 'experiment' as 'an operation carried out under controlled conditions in order to discover an unknown effect or law, to test or establish a hypothesis or to illustrate a known law'. An experiment, by definition, is a series of actions carried out as a means to achieving a consciously understood purpose. Experiments are done 'in order to' discover, test or establish something. If intentions are not part of the causal explanation of overt action then *experimentation* is illusory.

If the conscious will is epiphenomenal as Wegner describes it, then the scientist might *design* an experiment. He may consider a thesis and think up ways to test it. But on Wegner's view these conscious events have no causal or explanatory impact on overt actions. The scientist cannot *conduct* any experiments. Can we even say correctly that the scientist can engage in certain motions such as arranging blocks, talking to the subject and writing on a note pad? Without invoking intentions in an explanatory role, it is difficult to analyse the behaviours of

'arranging' versus random putting, 'talking' versus just making sounds, and 'writing' versus making scratchings.[2] But, however we describe these motions, on Wegner's analysis, it is not the case that the scientist engages in these movements 'in order to' achieve a conscious purpose. These actions are the direct result of, and fully explained by, unconscious causes, which causes may or may not have also caused the epiphenomenal thoughts. Thus the scientist's motions cannot, by definition, constitute experiments. If Wegner's conclusion is correct, no one has, or could, demonstrate experimentally that conscious will is an illusion.

Yet, when Wegner describes the actions he, his colleagues and his sources engage in he seems to assume that they perform experiments under this definition. So, for example, he writes (my italics for emphasis), 'Wegner and Wheatley (1999) conducted an experiment *to learn* whether people will feel they willfully performed an action' (2002, p. 74); '*In trying to understand* what kind of mental system would make a suppressed thought come back to mind again and again, my colleagues and I did a series of experiments' (p. 141); 'Lots of clever studies *have been aimed* at this distinction' (p. 179); 'in our experiments a variety of conditions were tested *to see whether* this tendency to answer correctly could be changed' (pp. 203–4); '*To ascertain whether* belief was indeed causal, on one study we directly manipulated belief' (p. 206). The implication is that, while the rest of us poor dupes are living an illusion when we suppose ourselves to be capable of intentional actions, Wegner and his colleagues can do what is impossible for us. They can act for reasons! Wegner allows that 'we' are all often seduced by the illusion, and so one might think that he includes himself and his colleagues (2002, pp. 341–2). But he does not go on to grant that no one has ever actually conducted experiments so there is no experimental evidence for his views. He seems to take it for granted that *qua scientist* he is above the rest of humankind, able to study our delusions but not sharing them. It is this attitude that is especially disquieting. Unless Wegner is actually superhuman, he should not exempt himself from the universal claims he makes. And if he cannot exempt himself, he should rescind the universal claims. The Golden Rule of experimental psychology regarding universal claims about human beings ought to be 'Attribute unto others only those characteristics you would have attributed unto you.'

The Christian tradition coupled freedom with human dignity and equality. Experimental psychology need not undermine that view. First and foremost, it should limit its conclusions to what the evidence warrants. Further, it can be conducted with respect and charity towards its

subjects and with the appropriate humility to recognize that conclusions which apply to the subjects of experiments, qua human beings, apply to the experimenters themselves. My prediction is that, while experimental psychology may uncover all sorts of interesting limitations regarding our ability to choose, it will not find that we are unfree in some philosophically interesting way. On the contrary, by apprizing us of our limitations, it can guide us towards taking more control over our choices and in that way help us to polish the reflection of the divine in ourselves – the very activity for which God gave us freedom in the first place.

Notes

1. Saint Augustine of Hippo is one of the chief architects of Christian philosophy and a consistent defender of the basic metaphysical and practical equality of human beings. He famously argued that government is a species of 'necessary evil' in that it is necessary due to sin, but it is disordered in that it gives some authority to use force against others. This is disordered because all human beings are fundamentally equal. And his antipathy towards some of the heresies he combated, including Pelagianism, stems at least in part from his disapproval of the claims of their adherents to be a moral elite superior to 'ordinary' human beings (Rogers, 2008, p. 129).
2. Perhaps a philosophical epiphenomenalist could manage it, but Wegner's view is both extreme and undeveloped. Note a related problem: his experiments often involve first-person reporting, but drawing conclusions about beliefs based on first-person reports seems justifiable only if there is some causal or explanatory connection between the beliefs and the overt action of making the sounds that constitute the reporting.

References

Burger, J. (2009) 'Replicating Milgram', *American Psychologist* 64: 1–11.
Dennett, D. (1984) *Elbow Room* (Cambridge, MA: MIT Press).
Doris, J. (2002) *Lack of Character* (Cambridge: Cambridge University Press).
Kane, R. (1996) *The Significance of Free Will* (Oxford: Oxford University Press).
Lewis, C. S. (1997) *Out of the Silent Planet, Perelandra, That Hideous Strength* (New York: The Quality Paperback Book Club).
Mele, A. (2009) *Effective Intentions* (Oxford: Oxford University Press).
Rogers, K. (2002) 'The Abolition of Sin', *Faith and Philosophy* 19: 69–84.
Rogers, K. (2007) 'Retribution, Forgiveness, and the Character Creation Theory of Punishment', *Social Theory and Practice* 33: 75–103.
Rogers, K. (2008) *Anselm on Freedom* (Oxford: Oxford University Press).
Skinner, B. F. (1976) *Walden Two* (New York: MacMillan Publishing Company).
Wegner, D. (2002) *The Illusion of Conscious Will* (Cambridge, MA: MIT Press).

12

The Compatibility of Science and Religion: Why the Warfare Thesis Is False

Michael Ruse

In the second half of the nineteenth century, a good number of scientists – prominent among whom was Darwin's great supporter Thomas Henry Huxley – subscribed to what is known as the 'warfare' thesis about the relationship between science and religion (Desmond, 1994, 1997). They argued that science and religion are in conflict, and that, if one holds to the one, one cannot hold to the other (Draper, 1875; White, 1896). The name of Galileo came up frequently in these discussions, and it was pointed out that it was only to be expected that the Catholic Church would have shown such opposition to so distinguished and fertile-thinking a scientist. Naturally enough, the recently published theory of evolution of the English naturalist Charles Darwin was another topic of conversation by these conflict theorists. Again it was suggested that one is faced with a stark dichotomy: either one accepts that organisms had natural origins or one accepts that we all arrived supernaturally. There is no other option.

Looking back, we can see now that this debate about the relationship between science and religion was as much social as intellectual. People like Huxley in Britain and his supporters both in that country and elsewhere (especially including the United States of America) were pushing strongly to reform society (Ruse, 2005b). They wanted a modern civilization, forward looking, and soundly based on science and technology. Rightly or wrongly, they saw the in-place churches as the major source of opposition, supporting as they did the establishment which wanted little or no change. The warfare thesis therefore was an important weapon of attack, whether or not it was fully sustainable either historically or theologically or philosophically.

At the beginning of the twenty-first century, we find that history is repeating itself. There is now a substantial fraction, drawn mainly from the sciences but also including supporters from philosophy and elsewhere, that is actively again promoting the warfare thesis. They argue that there is inevitable conflict and opposition between science and religion. Notable proponents of this position include the British science writer Richard Dawkins, author of the smash-hit bestseller *The God Delusion* (2007), as well as philosopher Daniel Dennett (2006), journalist Christopher Hitchens (2007), and sometimes student and neuroscientist Sam Harris (2004). As with the nineteenth century, one suspects that today the opposition is as much fuelled by social causes as by intellectual reasons. There is great tension about the rise and place of religion in society today. This applies both to the Christian religion, particularly the evangelical fundamentalist variety which is pushing not only to have its ideas included in science classrooms, but also a very conservative social agenda with respect to such things as abortion and homosexuality, and to other religions, particularly Islam. The horrific events in New York on September 11 2001 are never far from anyone's mind.

In this essay I shall say little or nothing about the social aspects of the warfare thesis. I have elsewhere discussed the nineteenth-century situation and no doubt others will discuss the twenty-first century better than I am able. I shall rather focus on the intellectual questions. In particular, I am interested in the question of whether or not science and religion are necessarily in conflict. The alternative position, one which incidentally has my support, is often known as the 'independence' position (Barbour, 1990). More recently, in a rather sneering manner, it has been referred to (particularly by the quartet mentioned in the last paragraph) as the 'accommodationist' position. As I set out, let me stress that I am not now trying to argue for the truth of any particular religion. I accept science but religion must find its own advocates. What I question here is simply whether how science is understood today poses an insuperable barrier to reasonable religious belief. Mainly because modern science has developed during the time of Christianity, I shall take that particular religion as my touchstone, although I trust that what I have to say could be extended to other religions.

My discussion will fall into two main parts. First I shall offer a general argument suggesting that the grasp of science is necessarily limited, and that questions reaching beyond that grasp are not only legitimate but fair game for the religious person to try to answer. The second part of the discussion will focus on a number of arguments that have been invoked to suggest that science and religion are necessarily incompatible.

1. Science and the Machine Metaphor

I start with the general case. This depends crucially on the fact that scientific understanding is inherently metaphorical (Kuhn, 1993). Scientific theories are far from just faithful descriptions of raw reality. From the beginning, they are attempts at understanding, at interpretation, and they always employ metaphors as they try to make sense of what we are experiencing, as they try to find ways to extend our grasp and look for new discoveries and connections. Take, for example, Darwin's theory of evolution through natural selection (Ruse, 2008). From the beginning, Darwin was thinking metaphorically. As he became an evolutionist, his thinking was done in terms of a 'tree of life' – a metaphor taken from the earliest chapters of Genesis. Darwin never thought of organic origins simply in their own terms, but instead as filtered through his cultural conception of life climbing ever-higher up the branches of a great diversifying tree. Then, as he started to move towards his key mechanism of natural selection, he immersed himself in the work of the animal and plant breeders of his day, thinking metaphorically as he pushed from his understanding in the human realm to what he thought might possibly be occurring in the natural order. The key insight for him came at the end of September 1838, when he was reading a work on population by the Revd Thomas Robert Malthus. Malthus was talking about population pressures and how there can never be food and space adequate to speak to them. Malthus therefore talked about the 'struggle for existence', something Darwin in a flash generalized to the whole world of animals and plants (Ruse, 1975).

From the beginning, Darwin recognized that this was metaphorical thinking. For a start, although sometimes there is a literal struggle, as often as not it is meant metaphorically as organisms compete (itself a metaphor), and moreover it is not really existence that we are talking about but reproduction. Then came the crucial move to natural selection. Darwin took the struggle and added in the ubiquitous variation we find in nature, arguing that there is an ongoing equivalent to the selective practices employed by animal and plant breeders. This 'natural selection' – as metaphorical as one can get – is the force of ongoing evolutionary change. But, as Darwin pointed out, it is not random change. It is rather change in the direction of adaptive advantage. Metaphorically, Darwin saw the organic world as if it had been designed – the eye is for seeing and the hand is for grasping – and selection speaks to this. In short, Darwin's theory is as metaphorical as a sermon or a piece of poetry (Ruse, 2005b).

What I want to suggest – and here, I emphasize, I am not being at all original but rather drawing on the consensus of those who write on the history of science – is that science today is governed by one overall, dominant metaphor, what is itself known metaphorically as a 'root' metaphor. This metaphor is that of a machine. Under the so-called 'mechanistic' viewpoint that dominates modern science, the physical world (and this today extends to the biological and even human world) is seen as if it is a giant machine (Hall, 1983). This metaphor goes back to the Scientific Revolution that occurred in the sixteenth and seventeenth centuries. The root metaphor for the Greeks, and for subsequent thinkers in the West down to the time of the Scientific Revolution, was that of an organism (Merchant, 1980). Plato, in particular, set the scene when, in his great dialogue the *Timaeus*, he likened the world itself to an organism, arguing that it has a soul of its own. (This, in medieval times, was known as the *anima mundi*.) It was for this reason that it was appropriate to think in terms of ends, of purposes, of what Plato's great pupil Aristotle was to call 'final causes'. With the Scientific Revolution all of this was swept away. The French philosopher and mathematician René Descartes argued that physical matter has no mind of its own (Garber, 1992). Its only distinguishing feature is that of extension. What we have, therefore, after the Revolution, is simply matter in motion going on invariably and inevitably, without end, according to blind laws. The British scientist-philosopher Robert Boyle (1996) was fond of likening the world to a clock. Of course you might object that machines, clocks in particular, have ends or purposes. But as historians of the Scientific Revolution point out, by the time the Revolution was over most scientists were ignoring the possibility of ends and thinking only of the ways in which the machinery works. In the words of one of the most distinguished commentators on the Revolution, God had become a 'retired engineer' (Dijksterhuis, 1961).

There is much more that one could say about the development of the machine metaphor in the four centuries since the Scientific Revolution. Most particularly, there was the coming of the machine metaphor to the biological world. It was here particularly that Charles Darwin's genius was important, as he showed that – as we have sketched above – the distinctive features of organisms could be explained in terms of blind law, no less adequately than the distinctive features of inorganic objects are explained in terms of blind law in physics and chemistry (Ruse, 1979). In the words of Richard Dawkins (1976, p. 22), organisms became 'survival machines'. Most recently, in the second half of the twentieth century, thanks to the coming of computers, students of humankind

have been making major advances in understanding how our thoughts and actions can be understood as the results of sophisticated electronic machines – machines otherwise known as computers. This of course is the basis of the new area of enquiry known as 'cognitive science'.

2. Unasked Questions, Unanswered Questions

I shall not dwell on details of the coming of the machine metaphor to science, but shall now move to explore its implications for our inquiry. The important point I want to make is that metaphors have both strengths and weaknesses, or if you do not want to talk in terms of 'weaknesses' then let us say 'limitations'. Thanks to metaphor, you can not only organize and explain what you have, but you can use it heuristically to explore new areas and problems. Think just in the area of biology, and see how regarding parts of organisms as functioning contrivances or machines has paid huge dividends. To take one example, consider the strange plates that run down the back of the dinosaur Stegosaurus. What are they for? One suggestion was that they might be for attack or defence, but it turns out that they are not very strongly made, so that suggestion almost certainly falls. Another suggestion is that they might be for sexual attraction, but since both males and females have the plates it is thought most unlikely that this is the case. The most recent suggestion is the plates serve the function of heat regulation, helping the cold blooded brute warm up in the sun in the morning and then to cool off in the heat of the day, thanks to the wind blowing across the plates (Farlow et al., 1976). This suggestion came because palaeontologists were aware of similar sorts of plates playing heat regulating roles in electrical generating stations. Because of a literal machine, we were able to understand the metaphorical machines that we find in dinosaurs.

The weaknesses or limitations of metaphors, as the late Thomas Kuhn (1962) pointed out, is that they succeed only by putting blinkers on us. They direct us to crucial areas, but at the same time rule certain questions out of court. In the terms of the metaphor, the questions simply do not make sense and should not be asked. Suppose, for instance, I say that my beloved is a rose. At once you know a number of things about her. You know that she is beautiful, at least in my eyes, that there is something fresh and young about her and that I am attracted to her. I suppose, if you know that I'm a little bit of a joker or a cynic, you might wonder perhaps if she is a little bit prickly. What you don't know is how good she is at mathematics. You don't know if she is strong

on the subject or weak. You don't know if she is bilingual. You don't know if she is English or French. It is not that these are not meaningful questions, they certainly are, but rather that my metaphor just does not address them at all. The same holds very much in science. On using a metaphor one focuses on some things, but not on others. Suppose one says that the DNA molecule incorporates a genetic code. This is a very useful metaphor, inasmuch as one is now led to ask questions about how to crack the code. But, if one asks other questions about the code, then they may well be inappropriate. One is not for instance asking about the religious affiliation of the molecules involved in the code. A question like this simply does not make sense.

Now to the root metaphor of the machine. I argue that there are certain questions that the metaphor simply rules out of court (Ruse, 2010). Under the metaphor, it is inappropriate to ask questions of this kind. I will list four such questions, although I do not want to claim that these are the only four questions that are ruled out. Moreover, I fully recognize that, in the future with a new understanding of machines and their power, one might think that some of the questions I focus on are indeed questions that can be answered in terms of machines. (In fact, as we shall see, already some would dispute my choice.)

The first question is about origins. One can ask how a machine was made, but ultimately where the materials came from is simply not part of the discourse. I can ask about where the copper used in the manufacture of the clock was dug up. But really where the copper came from is irrelevant to the making of the machine and its function. Translated to the metaphor of the world as a machine, I suggest therefore that science simply does not ask what is known as the 'fundamental question': Where did everything come from, or more generally why is there something rather than nothing? This question is simply not a scientific question in the sense that modern science just does not try to answer it at all.

The second question is about morality. Machines can be used for good, machines can be used for ill. But machines themselves are neither good nor bad. Take for instance an electric chair. Many people, myself included, would want to say there they can think of little good use for an electric chair. We disapprove of capital punishment under any circumstances. Many other people, people who would be indignant if it were suggested that they were not moral, would want to say that sometimes the electric chair can be used for good ends. If Hitler had survived the war, for example, then it would not only have been appropriate but

a good thing to execute him in the chair. Generalizing again about the world, the point is that the world-as-a-machine says nothing directly about morality. This is a point that philosophers have made many times in the past. As David Hume (1739) remarked, you cannot go from statements about matters of fact to statements about matters of obligation. It is fallacious. What I'm trying to do here is show precisely why such a move is fallacious.

The next question is probably more contentious. Machines are not conscious. This is the point made by the great German philosopher Leibniz (1714). Machines can do many things, but ultimately they are not conscious. So, science does not speak to the question, What is consciousness? Perhaps in this day and age one needs to modify this claim just a little bit. Possibly you might want to argue that, ultimately, really sophisticated computers will be conscious. Then, I would say that talking about the machine at the material level tells us nothing about why exactly the machines are consciousness, if by being conscious you mean sentience. There is a gap – what the philosopher-psychologist David Chalmers (1996) has called the 'hard problem' – between the way the machine works and the consciousness that emerges. So more generally what I'm saying is that, no matter what the triumphs of cognitive science, ultimately sentience goes unexplained. (As I said, this is more contentious. There are philosophers today who think that already cognitive science can explain sentience [Churchland, 1995; Dennett, 1992]. However, when you look at the arguments, there is inevitably a gap between the premises about the functioning of the brain or the computer and the conclusion about sentience. Along with many philosophical critics, I argue that there is a gap between the functioning of the material machine and consciousness.)

The final question is about ends. We have seen that, under the machine metaphor as used in today's science, questions about ends are ruled out. Science focuses only on the working of the machine and not upon the reasons why the machine exists in the first place. It will be remembered that God is a retired engineer. Hence, what I suggest is that modern science simply does not set out to answer questions about purposes. The Nobel prize-winning physicist Steven Weinberg (1992) has remarked that the more he does science the less he sees any purpose to the world. Not only am I not surprised, I would be very surprised if he did find purpose. The way we think about the world today using the machine metaphor does not allow for answers of this kind. Hence it is pointless to ask questions that demand such answers.

2.1. Answers?

We have four questions that science not only does not answer but that science does not even set out to answer. Why is there something rather than nothing? What is the ultimate basis of morality? What is sentience and its cause? What is the ultimate purpose of everything? I think it is possible to be a sceptic or agnostic on all of these questions. It's perfectly reasonable simply to say that one has no answers and cannot see any way of getting answers. Overall, this is my position. It is perfectly meaningful to ask why there is something rather than nothing. It is just that I don't have an answer! As far as morality is concerned, you can explain it away as an adaptation. I myself am inclined to say that this is enough and there is no need to go looking for foundations. But if you insist on foundations, then science is silent and so am I. Again, I am with those who not only do not have an answer to sentience but who are starting to think that perhaps there never will be an adequate answer (McGinn, 2000). Perhaps it is part of the way that our brains have evolved that this is a question that can never be resolved. This may seem like a bit of a cop out, but already in physics we are told that there are questions which we cannot answer. I am thinking about whether an electron is really a particle or really a wave. Heisenberg's Uncertainty Principle simply says that questions like these are off limits. There are no answers. Perhaps in cognitive science there is the equivalent to this kind of inexplicability. We know all about sentience except the most important questions, namely what it is and how it comes about. And certainly I have no answers to the ultimate purpose of reality. I don't see any purpose, so at this level I am with Steven Weinberg. Whether there is some ultimate purpose, however, seems to me simply beyond the grasp of science and once again beyond my grasp entirely also.

However, what I would argue is that my inability to offer answers does not preclude others trying to offer answers, so long as they do not offer scientific or quasi-scientific answers. You have to offer answers of a different kind. And here it seems to me is a place for the religious thinker. If the believer wants to offer answers, then I think it is perfectly legitimate for the believer to do so. I am not saying that the believer is beyond criticism, but I do not think that the believer can be criticized on the grounds of science. I am certainly saying that, if a believer insists that the Earth is 6000 years old, because this is what he or she learns in the Bible, then it is perfectly legitimate for the scientist to criticize such a claim. The believer is making a claim in the realm of science and as such is open to critique. I'm also saying that if the believer makes a claim that one wants to criticize on grounds other than science, it is legitimate to

make such a criticism. If, for instance, the believer makes a claim about a deity which seems internally contradictory or whatever it is legitimate to criticize. If, for instance, one argues that God is a necessary being, I think it legitimate for a philosopher to go after this claim, arguing that existence can never be necessary. (I am not now saying that the philosopher will be successful.) The point is that one is not now criticizing on the basis of science. And this is my position. Within these parameters – no trying to slip in science and recognize that you can (potentially) be faulted on other grounds – I would argue that the believer has every right to offer his or her solutions to the questions listed above.

Certainly it is the case that believers do offer answers to these questions. Staying – as restricted above – with Christianity, we find that there is a clear answer to the first question about why anything at all exists. The world is the creation of a god who did so as an act of love. God created out of nothing and He did so simply because He wanted to do so, and He wanted to do so because He saw that it would be better if He did so than if He did nothing. This is obviously not a scientific answer. It is a religious answer. I suggest – putting aside whether or not it is an adequate answer – in the light of the discussion above, it is certainly a legitimate answer.

What is the foundation of morality? Again the answer comes through loud and clear. It is that which God wants us to do. Morality is fitting in with the will of God. Notice that Christians realize that one has got to go a little beyond this (Quinn, 1978). One does not want to say simply that morality is God's will and that God might have done anything. Clearly, we do not want to say that God might have made killing a good thing. This sophisticated Christian response to this kind of worry is to point out that morality is bound up with the creation. God has made the world in a certain way and we are expected to behave in accordance with that way. In other words, we are expected to do that which is natural. So, for instance, wanton killing could never be right because it is unnatural. It is destroying God's creation without good reason. Morality therefore finds its foundation in God's will, but understanding God's will demands that we take due note of God's creation.

What is sentience? Again, note that we have got to be careful not to offer a scientific answer or an answer that is pretending to be religious but really is scientific. The Christian answer is that sentience is that which makes us 'in the image of God'. In some sense sentience or consciousness is that which puts us in tune with our creator. It is that which opens up the possibility of knowledge as well as of moral choice. Without these abilities we could not perform the role in the drama that

is taking place here on Earth. Sentience does not make us divine, but it opens up the future that we can share uniquely with the divine.

What is the purpose of it all? Ultimately, obviously, it is to share eternity with our creator. Exactly what this means is a matter of considerable debate among believers. But it is certainly intended to be a factual claim in some sense and not just a pious hope. The Earth was created to be the home of humans (and perhaps other organisms too), but in a sense this is only a temporary resting place. The Earth is a place of trial, in some sense, and depending on our performance here our future is determined accordingly. So the Earth does not have value just in itself but as a place of preparation for what is to come.

I want to stress again that I am not suggesting here that anybody has to accept any of these answers – or parallel answers that might be offered by members of other faiths. What I am arguing is that, given its inherent nature, modern science does not even attempt to answer the questions being asked. Therefore, it is legitimate for others to make the attempt.

3. Problems?

I have also stressed that I do not think anything said thus far makes religious answers immune to criticism. I wanted therefore now to change track somewhat and to look at three criticisms that have been made in the name of science, suggesting that the religious answers are not well taken. In other words, I want to look at criticisms that suggest even though it may be possible for the religious person (the Christian specifically) to offer a legitimate response in the light of science, the particular responses that the religious person wants to make are inadequate.

First let us raise the venerable problem of evil. Many Darwinians think that Darwin's theory of evolution poses an altogether-much-more-severe threat on this front than has been encountered hitherto. The traditional Christian approach to the problem divides the problem into two (Ruse, 2001). On the one hand, we have moral evil – human-caused, like Auschwitz – and on the other hand we have natural or physical evil – earthquakes, like the one that destroyed Lisbon. In both cases, Darwinism poses a threat. Moral evil is supposedly something introduced by the sin of Adam – God created us perfect – and Darwinism negates an original pair of humans. Physical evil is simply the result of laws and it is argued that the laws are a package deal, sometimes causing good things and sometimes bad things, but overall good. (Pain is uncomfortable but better pain than accidents where we are unaware that our bodies are in danger.) But natural selection, as we have seen, depends on the struggle

for existence. The struggle for existence is frequently very painful and so it seems almost capricious of God to have created using this mechanism. Would an all-loving, all-powerful God have chosen selection as his means of creation? Darwin himself worried about this issue. And Darwinians down to the present, especially including Richard Dawkins (1995), have made much of this point.

My aim is not to defend Christianity against the problem of evil. I am far from convinced myself that this is possible. Here, I am interested rather in whether or not the Darwinian critics are correct and whether science – Darwinian evolutionary theory in particular – poses an altogether-new dimension to the problem. I'm not sure that it does. Start with moral evil. Following St Augustine, it is argued that thanks to Adam we are tainted in some way and hence have an inclination to do bad things. Obviously, the evolutionist cannot accept Adam and Eve as literal, historical figures. However, the evolutionist does have an explanation of moral evil, namely that Darwinian selection is going to produce humans that are somewhat ambivalent mixes of good and ill. On the one hand, we are social beings and as such therefore have inclinations to work harmoniously with others. On the other hand we have to look after ourselves – in Dawkins's colourful metaphor, we have to be 'selfish genes' – otherwise we simply would not succeed in the struggle for existence. So what is argued by the Darwinian is that, rather than evil being the result of one person's bad act at some point in history, it is the natural outcome of a long, slow process of development.

Is this the end of the Christian story? It is certainly the end of the Augustinian story. But there are today Christian theologians who suggest that this Darwinian account is a more adequate *theological* answer to the problem of evil (Schneider, 2010). Rather than making a whole Christian story a result of one weak man's act of disobedience, and then the whole drama of Christ coming and dying on the cross for our sins being, as it were, a matter of playing catch up – what is often referred to semi-humorously as 'Plan B' – we see that, in a sense, God intended everything from the start. This is a tradition which goes back before Augustine to Irenaeus of Lyons, and many think it a more comfortable and reasonable interpretation of what it is to say that God had become to save us from our sinful nature. Because of our history we are but partial figures and the Incarnation and the Atonement were part of finishing the creation, not cleaning up a mess after the creation.

As far as physical evil is concerned, interestingly it may well be Richard Dawkins (1983) himself who comes to the Christian's rescue at

this point. He argues that the only way in which one can naturally get design or final cause – organic adaptation – is through natural selection. All other mechanisms are either false (for instance so-called Lamarckism, the inheritance of acquired characteristics) or inadequate (for instance saltationism, the idea of evolution through jumps). But if natural selection is the only way in which organisms could have been produced, and this includes human beings, then inasmuch as God decided to create through law natural evil had to follow. Remember it is never the claim that God can do the logically impossible. Only that he will always do that which is best. Of course, there is also the matter of whether or not God had to create through law. However, given the kind of beings that we are, had God not created through law then he would surely have been deceiving us – every part of our physical nature attests to our having evolved rather than having been created miraculously. So, it seems that there are strong reasons for thinking that natural evil or pain is not necessarily something that can be laid down as God's responsibility.

I turn next to a worry that many have about freedom. If one accepts the scientific perspective then one is committed to the belief that the world is governed according to universal law. The metaphor is that the world is like a gigantic clock, simply going through the motions. The clock obviously has no freedom of its own. If the hands point to 2 o'clock right now, and then in one hour's time the hands will point to 3 o'clock. The clock cannot of its own intent decide to speed things up and make the hands point to 4 o'clock or slow things down and make the hands point to 2.30. It is, however, an essential part of the Christian religion that humans are free agents. We have the power to choose between good and ill, even though regrettably all too often we choose the ill. God could presumably have made us all robots, doing nothing but good. However, in His wisdom, He saw that it is better that we be free even if sinful, than predetermined and unable to decide for ourselves. There seems therefore to be a major clash here between science and religion.

Moreover, there are those who think that particular sciences make matters even worse. In recent years, it has been claimed by critics that the attempt to apply evolutionary theory to human nature leads to a particularly pernicious form of determinism, often labelled 'genetic determinism' (Lewontin, 1991). It is argued that the human evolutionist – sometimes known as the 'human sociobiologist', although these days more likely as an 'evolutionary psychologist' – is committed to the view that everything we do and think is predetermined by the genes, as chosen by natural selection. Hence, for instance,

it is argued that the human evolutionist thinks that males are aggressors without any choice in the matter, simply because of their genes; whereas women are sexually coy, again simply without any choice.

That there is an issue here that needs resolving is beyond argument. However, there are fairly standard ways of answering the general complaint. In particular, many (quite independently of science) would argue that it is misleading to juxtapose the rule of law with free will. In fact there are those, notably the philosopher David Hume, who argue that only if we have the rule of law is freedom possible. Without law, then everything is purely random and this is not freedom. If for instance I commit a murder, but it is a purely random event, then I am not morally culpable. It is only when I'm seen against the background of my training and human nature and much more that one can properly prescribe praise and blame. This position, often known as compatibilism, argues that the real distinction must be drawn between being determined and being confined or in some way oppressed. A person in chains or hypnotized is not free. The person not so constrained, even though bound by law, is nevertheless free. So the general charge that science makes free will impossible is seen to fall to the ground.

What about the specific charge that human evolutionary biology introduces a new dimension of determinism, one which does indeed make free will impossible? Here, it is worth looking in a little more detail at the claims of the human evolutionist. It is indeed agreed that humans, like all other organisms, are the products of random variation brought on by natural selection. Hence, the genes that we have do play a vital causal role in our physical being and in our nature – both our thinking and our behaviours. But do note that human evolutionary biology today is a little more subtle than critics often appreciate. Most importantly, it is recognized that higher organisms like humans have rather different behavioural strategies than lower organisms like social insects. Ants, for instance, do everything by instinct. They do not require any teaching or prior experience. Hence, ants will follow a pheromone trail quite blindly to find food and then return to the nest fully laden. Given the success of ants in the world, this is obviously a very good adaptation because it leads to high reproductive success. However, working purely by instinct has its dangers and snares. In particular, if something goes wrong then it is very difficult to regroup and rethink and try a new strategy. Suppose an ant is out foraging and there is a tropical rainfall wiping away all of the chemical trails. It may well be that as a consequence literally hundreds of ants get lost and die, failing to return to the nest. The reason why the ant family can afford this kind of cost is that the queen

produces literally thousands of offspring, and the loss of even a large number is far from irreparable.

Humans have taken a somewhat different reproductive strategy. We only raise a few offspring and invest a great deal of parental care in them. This means that we cannot afford to lose our offspring as soon as some environmental event threatens our wellbeing. If every time there was a heavy rainfall we lost one or more of our offspring, we would very rapidly go extinct. Hence, humans have to have the ability to regroup and rethink when obstacles or adverse events come into our part. This we do, showing that at some level we have a degree of freedom not possessed by lower organisms like ants. We are still completely law-bound but we have something over and above just brute determinism. The philosopher Daniel Dennett (1984), for all that he is violently opposed to religious belief, has offered a good illustration of how humans work and a sense in which we can be said to have free will. He likens us to the Mars Rover, which when faced with rocks and other obstacles on the red planet does not give up but reassesses and moves around the obstacles. Everything is entirely law-bound but there is a dimension of freedom that the more simple machine would not have. Hence we can see that far from modern science standing in the way of Christian claims about free will, if anything it helps us to understand how such free will comes about.

4. Are Humans Necessary?

I want now to turn to a third problem, one which I believe is more challenging than the two just discussed. This is about the necessity of human existence. I do not think it is necessarily the case that the Christian must believe that we humans are uniquely the focus of God. If one takes seriously the comment in the Gospels about God knowing when every sparrow falls, it would seem that God has an intensive concern for all of His creation. But it is surely the case that humans do have a special place in God's heart, and that it was absolutely necessary that we humans appear. It might be impossible that we would not have been exactly as we are today. Perhaps we might have had blue skin, or maybe had twelve fingers. Possibly, although I'm not sure about this, it might have been possible for us not to have sexuality as we know it today. However, it is surely the case that the beings that were intelligent and with a moral dimension had to evolve. If this had not happened, then Christianity would be untrue.

Unfortunately, Darwinian evolutionary theory implies strongly that there is no necessity in the overall process. Stephen Jay Gould (1989)

used to say, if you play the tape of life several times, you will get different end results each time. Because natural selection is a relativistic process – this does not mean that it is tautological, but it does mean that what survives in one situation might not be what survives another situation – there is no absolute inevitability about what will come out at the end. And if this were not enough, it is always stressed by Darwinians that the basic building blocks of evolution – the mutations – are random, not in the sense of being uncaused but in the sense of not appearing according to the needs of their possessors. All of this seems to speak strongly against any claim is that human beings had to appear on the scene. Hence, it does seem finally that we have come to a point where modern science strikes strongly at the central claims of the Christian religion.

Those concerned to reconcile science and religion have not been unaware of this problem. One solution is simply to take the whole problem out of the realm of science. This is the position taken by the physicist-theologian Robert J. Russell (2008), who argues that perhaps God puts in direction down at the quantum level. Evolution is guided and humans will inevitably appear, even though we can never expect to see exactly how this works. However, my suspicion is that most feel rather uncomfortable with a solution like this. It is too close to a 'God-of-the-gaps' type solution, where when you don't know how to solve something you bring it in a miracle. It is a modern-day version of the kind of position endorsed by Darwin's great American supporter Asa Gray (1876), a position that Darwin totally repudiated. If God wants to create through unbroken law, then He should let the law get on with the job without constantly interfering in the process.

There are a number of attempts to solve the problem using just modern biology. Richard Dawkins (1986) argues that evolution frequently involves what are known as 'arms races'. Lines of organisms compete against each other and eventually better adaptations emerge. If one keeps this up long enough one can expect that large brains will emerge and so humans will have to appear. Interestingly, Darwin himself rather wrestled with this problem – although I do not think he was doing it in order to solve theological issues any more than that Dawkins – and he came up with a similar solution.

> If we look at the differentiation and specialisation of the several organs of each being when adult (and this will include the advancement of the brain for intellectual purposes) as the best standard of highness of organisation, natural selection clearly leads towards highness; for all physiologists admit that the specialisation of organs,

> inasmuch as they perform in this state their functions better, is an advantage to each being; and hence the accumulation of variations tending towards specialisation is within the scope of natural selection. (Darwin, 1959, p. 134)

While this is an ingenious solution, it surely does not guarantee that humans must appear. Suppose that the dinosaurs had not gone extinct 65 million years ago. Then it might well be that the continued existence of the dinosaurs would bar the rise of the mammals, and humans would still not have emerged – nor would there be any prospect of them ever emerging. So while it is the case that arms races may make intelligence somewhat more probable, it certainly does not offer the iron-clad guarantee that the Christian demands.

Recently Simon Conway Morris, a distinguished palaeontologist who is also a practising Christian, has been suggesting that the key item in solving this problem is the existence of ecological niches. He argues that certain liveable niches exist, independently of organisms. What happens is that organisms, powered by natural selection, push and strain until they find and enter into these niches. Thus, for instance, we find that there is a sabretooth tiger niche, which was entered independently by the placental mammals and by the marsupial mammals. Conway Morris suggests that there is obviously a cultural niche, one that we humans have occupied, and that therefore even had we not occupied it we might expect at some point that some organisms would so do.

> If brains can get big independently and provide a neural machine capable of handling a highly complex environment, then perhaps there are other parallels, other convergences that drive some groups towards complexity. Could the story of sensory perception be one clue that, given time, evolution will inevitably lead not only to the emergence of such properties as intelligence, but also to other complexities, such as, say, agriculture and culture, that we tend to regard as the prerogative of the human? We may be unique, but paradoxically those properties that define our uniqueness can still be inherent in the evolutionary process. In other words, if we humans had not evolved then something more-or-less identical would have emerged sooner or later. (Conway Morris, 2003, p. 196)

Again, one can see problems. First, many evolutionists are not convinced that niches do exist independently of organisms. These evolutionists argue that, to a great extent, organisms make niches

themselves. Hence, although we humans may live in a cultural niche, it is wrong to think that the cultural niche exists independently of our particular course of evolution. And in any case, the dinosaur problem still stands in the way. Had the dinosaurs not gone extinct, even if a cultural niche did exist, this surely is no hard guarantee that humans or any other mammals or organisms would have found their way into it. Again, therefore, we do not have enough for the Christian.

A third solution is to drop Darwinism almost entirely. This is the tack taken in a recent book by palaeontologist Daniel McShea and philosopher Robert Brandon. They introduce what they call the 'zero-force evolutionary law' or ZFEL: 'In any evolutionary system in which there is variation and heredity, in the absence of natural selection, other forces, and constraints acting on diversity or complexity, diversity and complexity will increase on average' (McShea and Brandon, 2010, p. 3). Obviously, if their law is well taken, then if there is enough time one would expect very complex organisms to evolve. The trouble is that there seems no guarantee that the complexity would be in any sense adaptive. More particularly, that it would be adaptive in the sense of producing large, functioning brains. McShea and Brandon quite openly acknowledge that they are following not so much in the footsteps of Charles Darwin, but in the footsteps of his contemporary Herbert Spencer. And this has always been the problem with Spencerian-type thinking. You may indeed get an end result, but whether it is the end result you want is another matter. Certainly, without more added, and the more almost inevitably points one in the direction of natural selection, one is going to end up with something which is not necessarily human-like. Again the Christian will remain dissatisfied.

My own feeling is that all of these solutions go about the problem in the wrong way. The demand about the inevitable appearance of humans is not so much a scientific demand but a religious demand. It is not that the Christian is saying what must happen on scientific grounds, but rather what is needed on theological grounds. For Christianity to work, humans must have appeared. I suggest therefore that in looking for a solution we should start with Christianity itself. Given a universe, there was no guarantee that on any particular playing of the Gouldian tape that humankind would occur. We could have many, many universes where either life did not get going at all or if it did get going did not lead to humans. However, because humans did appear by Darwinian processes, we know that they could have appeared by Darwinian processes. They needed a world in which this could happen. Our world, obviously. At some point in an infinite string of universes, either in time or space,

it would work. Now, it is a standard part of Christian theology that God is omnipotent. So, presumably God could have created all of those universes and at some point He would get what he wanted. Would not this be an awful drag on God's patience? Not at all! It is also part of Christian theology that God stands outside time and space. God is not waiting endlessly for things to happen. As Saint Augustine argued, for God the thought of creation, the act of creation, and the product of creation are as one. So there is no big problem there. Humans appear, and no special forces or explanations are needed above and beyond pure Darwinism.

Although today there are many physicists inclined this way, I am not arguing for this 'multiverse' position on scientific grounds, but purely on theological grounds. You might think that this is awfully wasteful. Why did God create only the universe that he knew would be the one where humans evolved? Here, I think we need to go back to quantum mechanics, and to the claim that there is no predictability about how individual events will occur. If we have God choosing only one universe out of many, then we are back to Robert J. Russell's position – God is guiding events. I want to argue that God simply did not get involved at this level, and hence he did not know or determine which way things would occur at the quantum level. This does rather imply that God at some level is limited, not knowing what will emerge. But this is no more a limitation on God's powers than that he cannot make $2 + 2 = 5$. But what about the waste? God is creating literally millions of universes which come to nothing. But, God does this already, because our universe is absolutely vast with what seem to be countless lifeless galaxies. Why did God not just confine his creative abilities to our solar system? In any case, who are we to presume that we know God's intentions and desires? Who is to say that God did not delight in the very act of creation and that the universe without human beings would be unpleasing or irrelevant to God's purposes? The Christian claim is that humans have a special lien on God's affection, not that we are exclusively the objects of God's concern and interest (Whewell, 2001).

5. Conclusion

My general position is that, because science is metaphorical, it is limited in the sense that there are certain questions that it does not even ask let alone answer. I argue also that it is therefore legitimate for others to attempt to offer answers, so long as they are not scientific. Here is a place where religion might function successfully. I am not arguing that one must be religious, and indeed suggest that one could be

entirely sceptical or agnostic. It is simply that if Christianity wants to try to answer the unanswered questions of science, then it is free to do so. I suggest also that some of the most obvious objections to this enterprise are really not well taken. Hence, overall I want to argue for the compatibility of science and religion, and suggest that one can legitimately be an accommodationist. More strongly, I would argue that the reasonable person should be an accommodationist.

References

Barbour, I. (1990) *Religion in an Age of Science* (New York: Harper and Row).

Boyle, R. (1996) *A Free Enquiry into the Vulgarly Received Notion of Nature*, ed. E. B. Davis and M. Hunter (Cambridge: Cambridge University Press).

Chalmers, D. J. (1996) *The Conscious Mind* (New York: Oxford University Press).

Churchland, P. M. (1995) *The Engine of Reason, The Seat of the Soul* (Cambridge, MA: MIT Press).

Conway Morris, S. (2003) *Life's Solution: Inevitable Humans in a Lonely Universe* (Cambridge: Cambridge University Press).

Darwin, C. (1959) *The Origin of Species by Charles Darwin: A Variorum Text*, ed. M. Peckham (Philadelphia, PA: University of Pennsylvania Press).

Dawkins, R. (1976) *The Selfish Gene* (Oxford: Oxford University Press).

Dawkins, R. (1983) 'Universal Darwinism', in D. S. Bendall (ed.), *Evolution from Molecules to Men* (Cambridge: Cambridge University Press), pp. 403–25.

Dawkins, R. (1986) *The Blind Watchmaker* (New York: Norton).

Dawkins, R. (1995) *A River Out of Eden* (New York: Basic Books).

Dawkins, R. (2007) *The God Delusion* (New York: Houghton, Mifflin, Harcourt).

Dennett, D. C. (1984) *Elbow Room: The Varieties of Free Will Worth Wanting* (Cambridge, MA: MIT Press).

Dennett, D. C. (1992) *Consciousness Explained* (New York: Pantheon).

Dennett, D. C. (2006) *Breaking the Spell: Religion as a Natural Phenomenon* (New York: Viking).

Desmond, A. (1994) *Huxley, the Devil's Disciple* (London: Michael Joseph).

Desmond, A. (1997) *Huxley, Evolution's High Priest* (London: Michael Joseph).

Dijksterhuis, E. J. (1961) *The Mechanization of the World Picture* (Oxford: Oxford University Press).

Draper, J. W. (1875) *History of the Conflict between Religion and Science* (New York: Appleton).

Farlow, J. O., C. V. Thompson and D. E. Rosner (1976) 'Plates of the Dinosaur Stegosaurus: Forced Convection Heat Loss Fins?', *Science* 192: 1123–25.

Garber, D. (1992) *Descartes' Metaphysical Physics* (Chicago: University of Chicago Press).

Gould, S. J. (1989) *Wonderful Life: The Burgess Shale and the Nature of History* (New York: W.W. Norton Co.).

Gray, A. (1876) *Darwiniana* (New York: D. Appleton).

Hall, A. R. (1983) *The Revolution in Science, 1500–1750* (London: Longman).

Harris, S. (2004) *The End of Faith: Religion, Terror, and the Future of Reason* (New York: Free Press).

Hitchens, C. (2007) *God Is Not Great: How Religion Poisons Everything* (New York: Hachette).
Hume, D. (1739) *A Treatise of Human Nature* (Oxford: Oxford University Press, 1940).
Kuhn, T. (1962) *The Structure of Scientific Revolutions* (Chicago: University of Chicago Press).
Kuhn, T. (1993) 'Metaphor in Science', in Andrew Ortony (ed.), *Metaphor and Thought*, 2nd edn (Cambridge: Cambridge University Press), pp. 533–42.
Leibniz, G. F. W. (1714) *Monadology and Other Philosophical Essays* (New York: Bobbs–Merrill).
Lewontin, R. C. (1991) *Biology as Ideology: The Doctrine of DNA* (Toronto: Anansi).
McGinn, C. (2000) *The Mysterious Flame: Conscious Minds In A Material World* (New York: Basic Books).
McShea, D., and R. Brandon (2010) *Biology's First Law: The Tendency for Diversity and Complexity to Increase in Evolutionary Systems* (Chicago: University of Chicago Press).
Merchant, C. (1980) *The Death of Nature: Women, Ecology, and the Scientific Revolution: A Feminist Reappraisal of the Scientific Revolution* (Scranton, PA: HarperCollins).
Quinn, P. L. (1978) *Divine Commands and Moral Requirements* (Oxford: Clarendon Press).
Ruse, M. (1975) 'Charles Darwin and Artificial Selection', *Journal of the History of Ideas* 36: 339–50.
Ruse, M. (1979) *The Darwinian Revolution: Science Red in Tooth and Claw*, 2nd edn (Chicago: University of Chicago Press, 1999).
Ruse, M. (2001) *Can a Darwinian Be a Christian? The Relationship between Science and Religion* (Cambridge: Cambridge University Press).
Ruse, M. (2005a) 'The Darwinian Revolution as Seen in 1979 and as Seen Twenty-Five Years Later in 2004', *Journal of the History of Biology* 38: 3–17.
Ruse, M. (2005b) *The Evolution–Creation Struggle* (Cambridge, MA: Harvard University Press).
Ruse, M. (2008) *Charles Darwin* (Oxford: Blackwell).
Ruse, M. (2010) *Science and Spirituality: Making Room for Faith in the Age of Science* (Cambridge: Cambridge University Press).
Russell, R. J. (2008) *Cosmology: From Alpha to Omega, the Creative Mutual Interaction of Theology and Science* (Minneapolis: Fortress Press).
Schneider, J. (2010) 'Recent Genetic Science and Christian Theology on Human Origins: An "Aesthetic Supralapsarianism" ', *Perspectives on Science and Christian Faith* 62: 196–212.
Weinberg, S. (1992) *Dreams of a Final Theory: The Search for the Fundamental Laws of Nature* (New York: Pantheon).
Whewell, W. (2001) *Of the Plurality of Worlds. A Facsimile of the First Edition of 1853: Plus Previously Unpublished Material Excised by the Author Just before the Book Went to Press; and Whewell's Dialogue Rebutting His Critics, Reprinted from the Second Edition* (Chicago: University of Chicago Press).
White, A. D. (1896) *History of the Warfare of Science with Theology in Christendom* (New York: Appleton).

Index

agency
  attributions of, 71–94, 173–4
agnosticism, 185, 186, 188, 233
Agurrie, Anthony
  arrow of time, 118
  eternal steady state inflation model, 112–24, 127
Almeida, Mike
  multiverse, 149, 154–5
altruism, 62
Ambjorn, Jan
  loopholes in Euclidean approach, 130–1
Anselm, 11, 27
Aquinas
  foreknowledge, 33, 36, 41
  heretics, 207–9, 221
  Plantinga, Alvin, 228
argument from evil; *see* problem of evil
Aristotle
  methodology of, 33–5
astrophysical cosmogony, 96
atemporalist position, 104, 107
atheism
  probability of, 171
  persistence of, 229–30
atheodicy, 50, 58
atomic theory, 54
Augustine
  Christian theology, 36, 254, 265
  creation, 137, 272
autonomy, 207, 243

Bach–Weyl theory of gravity, 101, 109, 110
Banks, Thomas
  contracting space, 116–17, 121, 136
Barrett, Justin
  prevalence of religion, 28, 32, 42, 176
belief polarization, 6, 225–32
Bible, 195–8, 262
big bang, 4, 96–7, 100–32
biological complexity, 3, 51, 64, 270
Bishop, John
  modest fideism, 231
black crunch, 117, 136
black hole, 100, 105, 106, 111, 116, 120–1, 136, 137, 143
Blackburn, Simon
  truth condition approach to modality, 14–18
Boethius, 36–42
Bojowald, Martin
  loop quantum gravity, 113–18, 123, 125, 127
Borde, Arvind
  singularity theorem, 108–9, 122, 123
Boyer, Pascal, 30–1, 34, 175
Boyle, Robert, 258
Brandon, Robert
  zero-force evolutionary law, 271
brute facts, 39
Bulbulia, Joseph
  value of theology, 30–1, 42
Burger, Jerry
  burger experiments, 247–8
Butler, Joseph, 61

Calvin, John, 208–9, 228
causation, 16, 55, 144, 178
  backwards, 17–18
  Hume, David, 16
Chalmers, David
  hard problem of consciousness, 261
character, 241–2, 247–8, 249, 252
cognitive
  process, 3, 71
  scientists, 2, 5, 28, 29, 31, 42, 43
  science of religion, 5, 28, 31, 32, 34, 43, 86, 89, 165–80
compatibilism, 267

compatibility of science and religion, 6–7, 237–75
conscious
  mind, 75–6
  will, 239–42, 252–3
consciousness, 14, 41–3, 75, 76, 78, 79, 105, 251, 261, 263
  modes of, 36–41
consequentialism, 207–9, 246, 248
  Contigentism, 147–8, 158
cosmological argument, 11, 53
  Kalam, 4, 95–6, 128, 129–33
cosmology, 2, 4, 96–128, 136, 145, 158
  quantum, 125–6, 129, 134
cost-benefit analysis, 213–15, 217–19, 221
creation
  creation *ex nihilo*, 125, 126, 128, 129, 135, 137, 144
  doctrine of, 99, 102, 133, 263
creationism, 3, 68
cyclic ekpyrotic model, 96, 135

dark energy, 118–19
Darwin, Charles, 3, 49–68, 69, 103, 255, 257, 265, 269–70, 271
  Darwinian argument from evil, 49–68
Darwinian, 63–7
Dawkins, Richard, 256, 258, 265, 269
demographics, 184
  explanations of, 191–2, 198
  of theism, 5, 184–98
Dennett, Daniel, 4, 31, 34, 42, 71–3, 242, 256, 261, 268
  intentional stance, 4, 71–3, 81–6, 89
Descartes, Rene, 67, 258
determinism, 191–2, 266–9
deus deceptor argument, 177–80
divine attributes
  omnibenevolence, 2, 53, 56, 58, 153, 178
  omnipotence, 2, 49, 53, 55, 56, 57, 58, 64, 65, 172, 272
  omniscience, 53, 55, 56, 57, 58, 64, 82
  timelessness, 2, 22, 28–43, 95, 144
  necessary existence, 2, 11–28
divine hiddenness, 178, 191, 229
doctrine of creation, 99, 102, 133, 263
Donaldson, Margaret
  modes of consciousness, 37–41
Doris, John
  lack of character, 239–53
Draper, Paul
  duplicate universes, 144–5
  God as benefactor, 148
  restrictions on universe, 152, 157
  science and religion, conflict of, 255
dual process theory of mind, 75, 89

Ellis, George
  contracting universe, 111–14
entropy, 115–21, 125–7, 128, 136, 137
epiphenomenalism, 241
epistemological, 12, 15, 232
  naturalism, 105
Everett, Hugh, 126, 143
  Quantum mechanics, many worlds interpretation of, 143
evolution
  natural selection, 3, 49, 51–68, 177, 257, 264, 266, 267, 269–71
  theory of, 49–71, 79, 103, 71–5, 78, 85–8, 255, 257, 265, 266, 269, 270, 271
  of space-time, 130
  of consciousness, 79
evolutionary
  analysis, 74
  biology, 3, 49–50, 267
  psychology, 3, 71, 74, 88, 175
exclusivism
  salvific, 6, 205–10, 217–20
  interventionist salvific, 210–20, 221
experience-enhancing theology, 3, 36–43
experimental philosophy, 185

Fall, the, 230
Feynman, Richard
  cosmology, 109, 134
fine-tuning, 60, 111–15, 122, 143, 151, 173
Fischler, Willy
  contracting space, 116–17
Fodor, Jerry, 75–7

folk
  physics, 77–82, 90
  psychology, 77, 80, 81
foreknowledge, 33, 35, 40
formation of character, 237–9
free will, 68, 193, 197, 199, 200, 230, 237, 239, 242, 267–8

Galileo, 255
general relativity, 96, 100, 104–9, 112, 113, 115, 119, 121, 129–31, 134, 136
genetic fallacy, 167–8
geodesics, 108–9, 123, 128
Geroch, Robert
  singularity, 100
God
  attributes; *see* divine attributes
  being, necessary, 11–28, 263
  concept of, 1, 12, 21, 177
  goodness, source of, 21, 23
  of-the-gaps, 97, 168, 169–70
Gould, Stephen Jay, 268
Gratton, Steve
  arrow of time, 118
  eternal steady state inflation model, 112–24, 127
Gray, Asa
  Darwin, defender of, 53, 55, 269
Gregersen, Neils Henrik, 175
Guth, Alan
  singularity theorem, 108–9, 122, 123

Hawking, Stephen, 109, 127, 129, 130, 131, 132
  Hawking–Bekenstein entropy law, 120
  Hawking–Penrose singularity theorem, 107
Heisenberg uncertainty principle, 132, 262
Helm, Paul, 32
heretics, 207–8
heuristic concept, 73, 77, 79, 80, 81, 83
Hudson, H., 144, 145, 148, 154, 159
  plentitudinous hyperspace, 153, 160
Hume, David, 15–16, 26, 261, 267
  causation, 16
Huxley, Thomas Henry, 255

incompatibility of science and religion, 256
independence position, 7, 256
infinite contraction, 111
instrumentalism, 72–3, 85
intelligent designer, 3, 143
inverted god-of-the gaps objection, 168, 169–70
Irenaeus, 265

James, William
  pragmatism, 72, 73
  origin of belief, 181
Jedi, 187–8
Johnson, Samuel, 39
Jurkiewicz, Jerzy
  loopholes in Euclidean approach, 130–1

Kalam cosmological argument, 4, 95–6, 128, 129–33
Kane, Robert, 240
Kenny, Anthony
  foreknowledge, 33–5
Kuhn, Thomas
  science and metaphors, 257, 259

Laplace's demon, 82, 84, 89
Leibniz, Gottfried
  best possible world, 146
  conscious machines, 261
Lewis, C. S., 243
Lewis, David, 145, 158, 207, 209
libertarians, 192, 242
Libet, Benjamin, 240
Linde, A. D.
  eternal inflation view, 143
Loll, Renate
  loopholes in Euclidean approach, 130–1

machine metaphor, 257–61, 265–6
Maitzen, Stephen
  argument from evil, 190–4

Manson, Neil
  precautionary principle, criticism of, 216, 219
Marenbon, John
  Augustine, 36
McGinn, Colin
  mysterianism, 231–2
McHarry, J. D.
  identity of indiscernibles, 145
  unsurpassable universe, 147
  dissolving the problem of evil, 152
McShea, Daniel
  zero-force evolutionary law, 271
Mele, Alfred
  free will, 237, 240
Milgram experiments, 244–8
Mithen, Steven
  religiosity as feature of society, 166–7
monotheism, 71, 87, 184
Monton, Bradley
  big bang, 96
  duplicate universes, 145
  multiverse, 148–50, 155, 156
  problem of omnibenevolent being, 153–4
moral
  agents, 67
  character, 61, 67
  evil, 238, 264, 265
  standing, 87–9
morality, 172, 206–7, 260–3
Morris, Simon Conway
  ecological niches, 270
Muller–Lyer illusion, 75–6, 78, 80
multiverse, 4, 106–8, 124, 126, 137, 143–58, 160, 272
  theistic, 143–59
  scientific, 143–51
  timeless, 106–7, 113
  contigentism, 147–8
  level IV, 143
mysterianism
  McGinn, Colin, 231–2
mystical experiences, 1, 42

natural selection, 3, 49, 51–68, 177, 257, 264, 266, 267, 269–71
natural systems, 85, 87, 88
natural theology, 52–3, 97, 99, 231
naturalism, 54, 184
naturalistic explanations of religion, 165–80
necessitarianism, 146–8
necessity
  de dicto/*de re*, 12–14
  projectivist account of, 12, 15–19
  metaphysical, 15, 17
  divine, 11–15, 21–6
  logical, 18, 26
  of humans, 268–72
no logical space arguments, 166
no-design hypothesis, 57–9, 63–6, 68, 69

O'Connor, Timothy
  objective threshold, 145, 147–9
  argument from evil, 154–7
Ockham's razor, 150
omnibenevolence, 2, 53, 56, 58, 153, 178
omnipotence, 2, 49, 53, 55, 56, 57, 58, 64, 65, 172, 272
omniscience, 53, 55, 56, 57, 58, 64, 82
ontological argument, 13–15
Oppy, Graham
  universe, beginning of, 96
  argument from evil, 156–7

Paley, William, 3, 52, 58
Parikh, Maulik
  antipodal observers, 122, 123
Parmenides, 100
Penrose, Roger, 111, 119–21, 127–8, 130
  cyclic conformal cosmogony, 109–10
  Hawking–Penrose singularity theorem, 107
Perkins, R. K., 154–6, 160
  persistence of, 229–30
personhood, 87–9
philosophy
  of mind, 3, 71, 88
  of religion, 4–6, 32, 143, 151, 158, 165, 166, 180, 225
physical evil, 264–5

physical reality, 4, 104, 105
physics, 4, 83, 96, 97, 99, 103, 105, 109, 110, 112, 120, 145, 158, 258, 262
  folk, 77–82, 90
  laws of, 84, 89, 90
  quantum, 56, 121
Pike, Nelson
  timeless God, 32
Pitts, Brian J., 96–103, 108–10, 115, 128, 133
  cosmic destroyer argument, 100, 133
Plantinga, Alvin, 181, 228
possible worlds, 14, 25, 102, 110, 133, 146, 147, 150–3, 158
pragmatism, 72–3, 85
precautionary principle, 213, 215–16, 218–20
problem of evil, 3, 4, 49–58, 143, 151–8, 160, 190, 264, 265
  Darwinian argument from evil, 49–68
projectivist; *see* necessity
counterfactual thinking, 12–28
psychologists, 5, 67, 74, 75, 237, 239, 242, 244, 245, 246, 248, 249
psychology, 6, 12, 26, 37, 75, 238
  developmental, 2, 37
  evolutionary, 3, 71, 74, 88, 175
  experimental, 7, 237–54
  of counterfactual thinking, 12–28
  folk, 77, 80, 81
Pyysiäinen, Ikka
  model of creation of theism, 34, 36, 42

quantum
  cosmology, 125, 126, 134
  fluctuations, 113, 115, 116, 125, 131
  geometrodynamics, 126, 130
  gravity, 103–6, 108, 109, 113, 114, 117, 125, 127, 128, 130–2, 134, 136
  physics, 56, 121

Quine, W. V., 12–13, 26, 79, 105
  epistemology, naturalised, 105
  modal involvement, three grades of, 12–15

radiation, 116, 119, 120, 136
rationality of theism, 174, 175–7
Rea, Michael
  epistemological naturalism, 105
  hyperspace criticism, 153
reformed epistemology, 176
religious
  agency, 71, 85–6, 88
  belief, 3, 6, 28, 32, 42, 71, 73, 82, 85, 86, 165–81
  tolerance, 2, 6, 209–10
reproductive success, 60, 73–4, 85–9, 267–8
Rowe, William
  evidential problem of evil, 224
Russell, Robert J.
  direction at quantum level, 269

salvation, 6, 191, 193–5, 197–8, 199, 205–21, 238
Schellenberg, John, 168, 178, 199, 229
  divine hiddenness, 178, 199, 229
  supernaturalism, 168
Schrodinger, Erwin
  de Sitter space, 122
scientific revolution, 258
Scotus, Duns
  forced baptism, 208–9
self-knowledge, 43, 198
Sellars, Wilfred, 81, 90
  manifest/scientific imagine, 81, 90
sensus divinitatis, 175–7, 178, 180, 228, 230
sin, 180, 207, 231, 238, 254, 264
singularity, 111, 112, 115, 117–20, 124, 127, 128, 129, 133, 135
Skinner, B. F., 243
Smith, Quentin
  criticisms of Hawking, 127–8, 130, 133
Smolin, Lee, 106, 131, 143
  fecund universe theory, 143
space-time, 4, 96–101, 104, 106, 110, 119–20, 122–3, 127–33

Sperber, Dan
representation, epidemiology of, 170
spiritual stance, 4, 72, 85–9
Steinhardt, P. and Turok, N.
cyclic model, 143
Stewart, M.
distinguishing universes, 145
suffering, 24, 53–61, 66–8, 153–4, 156–7, 190, 206, 250
supernatural, 1, 28, 29, 30, 170, 171, 191, 194, 231, 255
agents, 167–70
belief, 177, 184
entity, 96
god, 177
realm, 167
supernaturalism, 168
Swinburne, Richard, 65
God's existence, argument for, 224–5

Talbott, Tom
theistic multiverse, 149
technical theology, 3, 35–6, 43
Tegmark, Mark
level IV multiverse, 143
theism
classical, 186–8
Darwinian, 63–7
demographics of, 5, 184–98
singularity argument for, 96
knowledge of, 189
multiverse-, 157–8
open, 35
orthodox, 50, 53, 55–6, 58–9
rationality of, 174, 175–80
theistic evolutionists, 63–4
theistic metaphysics, 67
theodicy, 67, 68, 233
theology, 1, 9, 26, 28–43, 144, 175, 176
technical, 3, 35–6, 43
theory of everything, 112
time's arrow, 121
two system hypothesis, 75

universalism, 195–8

Vaas, Rudiger, 113, 115
van Inwagen, Peter, 149
Vilenkin, Alexander, 108–9, 112, 116, 122, 124, 129, 131–2, 135, 137
tunnelling from nothing approach, 109

Ward, Keith
Donaldson, Margaret, 38, 40
warfare thesis, 7, 255–73
Wegner, Daniel
The Illusion of Conscious Will, 239–53
Weinberg, Steven
purpose in the world, 262
Weyl Curvature Hypothesis, 119–21
Wheeler–DeWitt
approach to quantum gravity, 104, 106, 125, 126,
Whitehouse, Harvey
idea of religion, 30, 31, 34, 42

Zeh, H. D.
double big bang, 125–6
zero-force evolutionary law, 271
McShea, Daniel, and Brandon, Robert, 271